GFZ

Helmholtz Centre
POTSDAM

HELMHOLTZ CENTRE POTSDAM
GFZ GERMAN RESEARCH CENTRE
FOR GEOSCIENCES

Sebastian Specht, Oliver Heidbach
Fabrice Cotton, Arno Zang

Data-driven earthquake focal mechanism cluster analysis

Scientific Technical Report STR17/01

www.gfz-potsdam.de

HELMHOLTZ
ASSOCIATION

Recommended citation:

Specht, S., Heidbach, O., Cotton, F., Zang, A. (2017): Data-driven earthquake focal mechanism cluster analysis, (Scientific Technical Report STR ; 17/01), Potsdam: GFZ German Research Centre for Geosciences.
DOI: http://doi.org/10.2312/GFZ.b103-17012

Imprint

HELMHOLTZ CENTRE POTSDAM
GFZ GERMAN RESEARCH CENTRE FOR GEOSCIENCES

Telegrafenberg
D-14473 Potsdam

Published in Potsdam, Germany
May 2017

ISSN 2190-7110

DOI: http://doi.org/10.2312/GFZ.b103-17012
URN: urn:nbn:de:kobv:b103-17012

This work is published in the GFZ series
Scientific Technical Report (STR)
and electronically available at GFZ website
www.gfz-potsdam.de

Sebastian Specht, Oliver Heidbach
Fabrice Cotton, Arno Zang

Data-driven earthquake focal mechanism cluster analysis

Scientific Technical Report STR17/01

Data-driven earthquake focal mechanism cluster analysis

Sebastian Specht[1,2]
Oliver Heidbach[2]
Fabrice Cotton[1,2]
Arno Zang[1,2]

1 Institute of Earth and Environmental Science, University of Potsdam, Potsdam, Germany
2 Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Potsdam, Germany

Contents

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

List of Figures

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Abstract

Earthquake focal mechanism solutions (FMS) form the basic data input for many applications, e.g. stress tensor inversion or ground-motion prediction equation estimation. In these applications the FMS data is usually binned spatially or in predetermined ranges of rake and dip based on expert elicitation. However, due to the significant increase of FMS data in the past decade an objective data-driven cluster analysis is now possible. Here we present the method ACE (Angular Classification with Expectation-Maximization) that identifies clusters of FMS without a priori information. The identified clusters can be used for the classification of the Style-of-Faulting and as weights for FMS data binning in the aforementioned applications. As an application example we use ACE to identify FMS clusters according to their Style-of-Faulting that are related to certain earthquake types (e.g. subduction interface) in northern Chile, the Nazca Plate and in Kyūshū (Japan). We use the resulting clusters and weights as a priori information for a stress tensor inversion for these regions and show that uncertainties of the stress tensor estimates are reduced significantly.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Chapter 1

Introduction

Earthquake focal mechanism solutions (FMS) are of key importance for understanding rupture kinematics and the geometry of faults at depth. They are used e.g. to derive the orientation of maximum horizontal stress and the stress regime for the World Stress Map where they are the key contributor to this global data base (*Zoback*, 1992; *Heidbach et al.*, 2010). Furthermore, knowledge of the Style-of-Faulting (SoF) and the fault plane orientation are crucial in several fields of seismology, e.g. seismic hazard (GMPE, b-value) and in studying the Earth's stress field and changes within it (*Meier et al.*, 2014). Furthermore, data sets of FMS are used to investigate the seismotectonic setting of sub-regions and to derive the stress tensor orientation and relative magnitudes by means of a formal stress inversion (e.g. *Gephart and Forsyth*, 1984; *Michael*, 1984). In these applications the FMS data are usually binned spatially or in predetermined ranges of rake and dip based on expert assessment.

Several methods have been suggested to investigate clusters of FMS and moment tensors (MT), respectively. FMS represent the double couple component of a rupture plane, while moment tensor solutions provide also non-double-couple components. Due to the nodal plane ambiguity of the double couple, cluster analyses so far published are limited to certain aspects of FMS. E.g. *Frohlich* (1992) reduces the FMS (MT) to the plunge of the vertical components of the PBT axes. By expressing the focal mechanism in terms of their PBT axes, the nodal plane ambiguity is avoided. An advantage of this technique is the way data are presented: The complex geometry of the FMS can be represented in a two-dimensional ternary diagram, where each triangle corner is one end member of the Andersonian Style-of-Faulting (SoF), i.e. normal, reverse and strike-slip faulting. FMS in the center of the plot are classified as "oblique" or "odd". Event clusters are e.g. investigated for their consistency, a metric based on the scalar moment (*Frohlich and Apperson*, 1992).

Kagan (1991) defines the relative similarity of the FMS (MT) based on the shortest angle to rotate one FMS into another. A major advantage is that results are given in one parameter only, the minimum rotation angle between the FMS (frequently called Kagan angle). The distribution of the rotation angle follows a wrapped Cauchy distribution (*Ka-*

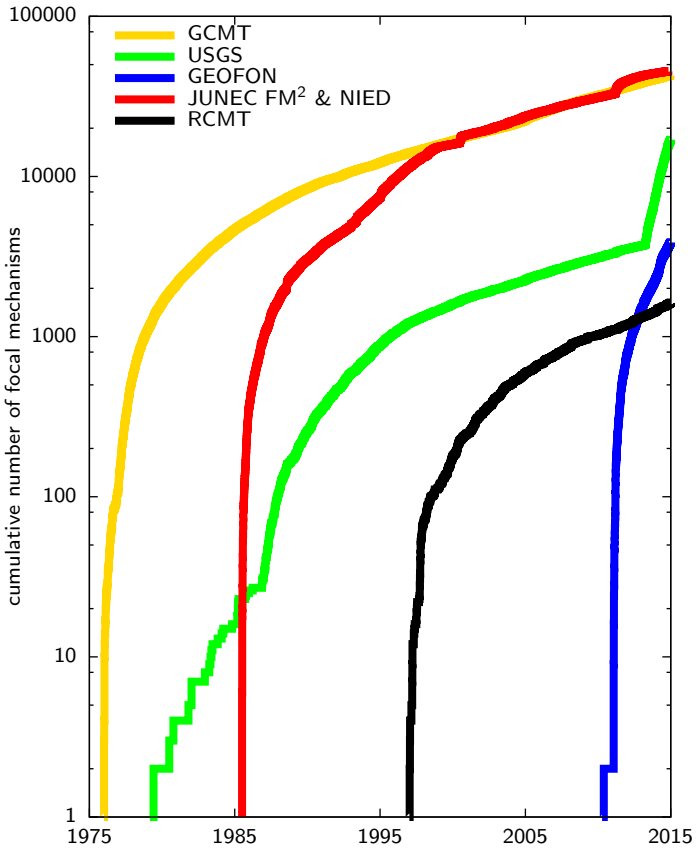


Figure 1.1: Catalog growth over time of GCMT (Global Centroid-Moment-Tensor, *Dziewonski et al.* (1981); *Ekström et al.* (2012)), GEOFON (GEOFosrchungsNetz, *Hanka and Kind* (1994)), USGS (United States Geological Survey), JUNEC FM^2 (Japan University Seismic Network Earthquake Catalog of First-Motion Focal Mechanisms, *Ishibe et al.* (2014)), NIED (National Research Institue for Earth Science and Disaster Prevention, Japan) and RCMT (European-Mediterranean Regional Centroid Moment Tensor Catalog, *Pondrelli et al.* (2011)). GCMT, GEOFON and USGS are global catalogs, while the others are regional catalogs: JUNEC FM^2 & NIED and RCMT.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

gan, 1992). Another cluster analysis method is the so-called density-based clustering. It is based on the DBSCAN algorithm of *Ester et al.* (1996) and assumes that FMS (MT) are distributed with two different "densities" of events (*Cesca et al.*, 2014). This density can be measured in different ways, in case of FMS/MT usually some angular metric, e.g. the above mentioned Kagan angle. Events with high densities (short distance metric values) are joined to one cluster; events with low densities (large distances) are considered to be noise and not included in a cluster. The advantage of this method is that no assumption on the underlying distribution of the data is required. The clusters are defined by two parameters, the minimum number of data per cluster and the threshold of the metric length at which a cluster is identified. Thus, the definition of "high" and "low" densities as well as the minimum size of a cluster are user based.

Another common way of cluster assignment thresholds are based on the expertise and experience of scientists by including additional information such as focal depth and information from structural geology. Though straightforward in its implementation, the final choice of the selected data lacks transparency and does not provide a measure for the quality of the choice. Furthermore, two of the above mentioned techniques are also to some extend subject to expert based judgment. These are the event class boundaries in the ternary plot representation and the two cluster defining parameters of the DBSCAN based method.

In order to avoid the a priori expert assessment of FMS data sets an objective data-driven cluster analysis is needed. This analysis should also deliver weights in terms of the quality of the identified clusters. Given that the amount of FMS data has increased significantly in the past decade (Fig. 1.1), a purely data-driven earthquake focal mechanism cluster analysis is now possible. We develop in this paper the method ACE (Angular Classification with Expectation-Maximization) which identifies clusters of FMS without a priori information. The physical fundamentals of our FMS cluster algorithm are based on elasticity theory using the 3D Cauchy stress tensor to formally describe the stress state at a point in a 3D volume. The algorithm is not a stress tensor inversion, but it also differs from the purely descriptive clustering methods presented e.g. in *Frohlich and Apperson* (1992) or *Cesca et al.* (2014). ACE follows basic assumptions about the distribution of random stresses (*Kagan*, 1990). It is a purely data-driven method that investigates the full scope of FMS data by considering both nodal planes and as little assumptions as possible. All parameters controlling the results are estimated in the process, thus mitigating effects from expert based judgment.

The identified clusters can be e.g. used for the classification of the Style-of-Faulting and as weights for FMS data binning in the aforementioned applications. We show how the results of ACE can be used to improve the binning of FMS data for the stress tensor inversions. The improvement is achieved by separating the FMS data into more consistent subpopulations of FMS with weights. The examples from northern Chile (subduction interface), the Nazca Plate and in Kyūshū (Japan) exemplify that the uncertainties of the stress tensor orientation resulting from a formal stress inversion are reduced significantly.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Chapter 2

Cluster analysis of focal mechanisms

2.1 Theory

2.1.1 Stress definition

An earthquake can be seen as a sudden deformation process resulting in slip caused by the ambient stress field in the surrounding rock. The 3D stress state is described by a symmetric 3×3 tensor $\boldsymbol{\sigma}$ (*Zang and Stephansson*, 2010; *Jaeger*, 1979).

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix} \tag{2.1}$$

The stress $\boldsymbol{\sigma}$ can be decomposed into four terms (*Kagan*, 1990, and ref. therein):

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_l + \boldsymbol{\sigma}_t + \boldsymbol{\sigma}_f + \boldsymbol{\sigma}_r, \tag{2.2}$$

where $\boldsymbol{\sigma}_l$ is the lithostatic stress, $\boldsymbol{\sigma}_t$ the regional tectonic stress, $\boldsymbol{\sigma}_f$ the stress related to prior earthquakes, and a random stress term $\boldsymbol{\sigma}_r$. The first three terms are assumed to be related to macroscopic phenomena and therefore constant and non-random. The random stress term is due to defects in the rock mass.

Kagan (1990) showed that the orientation of the random stress $\boldsymbol{\sigma}_r$ follows a Cauchy distribution. Based on this assumption, the distribution of the focal mechanism orientations follows a wrapped Cauchy distribution (the concept of wrapped distributions is introduced in the next section) (*Kagan*, 1990, 1992).

We use the assumption of stress decomposition in Eq. 2.2 to derive a statistical model to describe the distribution of the three FMS angles, namely strike, rake and dip.

In the following we derive the relationship between stress $\boldsymbol{\sigma}$ and the slip and normal vectors of a focal mechanism and their associated angles.

The three eigenvectors of $\boldsymbol{\sigma}$ indicate the principal stress directions and the eigenvalues the principal stress magnitudes. We define the principal stress orientations as $\hat{\mathbf{v}}_1$, $\hat{\mathbf{v}}_2$, $\hat{\mathbf{v}}_3$ with associated stress magnitudes as $S_1 \geq S_2 \geq S_3$ and compressive stress to be positive. The traction $\mathbf{T}$ is the force resulting from $\boldsymbol{\sigma}$ acting on a surface with normal $\mathbf{n}$ (Fig. 2.1).

$$\mathbf{T} = \boldsymbol{\sigma}\mathbf{n} \tag{2.3}$$

The normal stress σ_n follows from

$$\sigma_n = \mathbf{T} \cdot \mathbf{n} \tag{2.4}$$

and the vector of maximum shear stress $\mathbf{s}$ is

$$\mathbf{s} = \mathbf{T} - (\mathbf{T} \cdot \mathbf{n})\mathbf{n} \tag{2.5}$$

The slip $\hat{\mathbf{d}}$ on a fault is oriented in the direction of maximum shear stress (*Wallace*, 1951). In case of a preexisting fault, *Bott* (1959) suggested that fracturing occurs in the plane in which the strength was first exceeded and the direction of the initial slip is defined by the direction of the maximum shear stress (Fig. 2.2). Furthermore it is assumed, that fault planes are planar and that stress perturbations and rotations on the fault surfaces are neglected. Under these conditions, it follows, that the slip direction $\hat{\mathbf{d}}$ and the normal vector of the rupture plane $\hat{\mathbf{n}}$ are located in the plane spanned by $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_3$, or in other words, with a normal vector parallel to $\hat{\mathbf{v}}_2$. This implies that $\hat{\mathbf{v}}_2 = \hat{\mathbf{b}} = \hat{\mathbf{n}} \times \hat{\mathbf{d}}$. *Lisle* (2013) pointed out that directions up to 26° off the maximum shear orientation are at still at 90 % of the shear stress magnitude. Thus, slip and shear stress orientations are not necessarily well aligned. However, the applicability of the Wallace-Bott hypothesis to earthquake rupture processes has been shown by several studies (e.g. *Dupin et al.*, 1993; *Pascal*, 2002).

2.1.2 Representation of a Focal Mechanism Solution

Expressing a auxiliary/rupture plane of an earthquake with normal vector $\hat{\mathbf{n}}$, slip vector $\hat{\mathbf{d}}$ and the cross product $\hat{\mathbf{b}} = \hat{\mathbf{n}} \times \hat{\mathbf{d}}$ in terms of the strike (ϕ), rake (λ) and dip (δ)

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

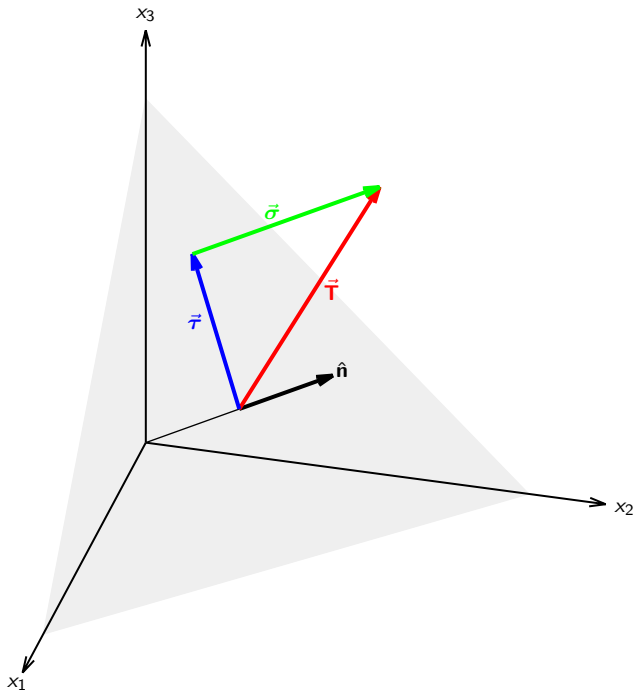


Figure 2.1: Stress vector $\mathbf{T}$ (red) acting on a plane (gray area) in 3D space. The stress vector can be decomposed into its normal (green) and shear components (blue). The normal stress component is orthogonal to the plane (parallel to $\hat{\mathbf{n}}$). Its magnitude is given by Eq. 2.4. The shear stress lies in the plane (parallel to it) with orientation given in Eq. 2.5. The eigenvectors of the stress tensor $\boldsymbol{\sigma}$ are orientated parallel to the three axes x_1, x_2, x_3.

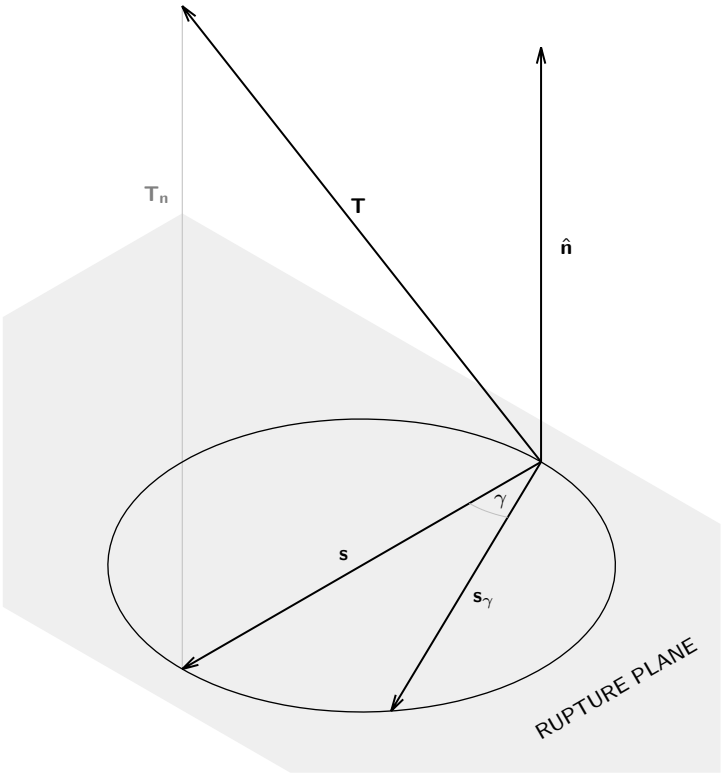


Figure 2.2: Orientation and magnitude of the resultant shear stress / slip on a plane. Maximum shear stress is orientated parallel to stress vector projected onto the rupture plane (Eq. 2.5). According to the Wallace-Bott hypothesis slip is orientated in maximum shear orientation. However, even directions 26° off the maximum shear orientation are still at 90 % of the magnitude of the maximum shear stress.

angles (*Kanamori and Cipar*, 1974):

$$\hat{\mathbf{n}} = \begin{pmatrix} -\sin\delta\sin\phi \\ -\sin\delta\cos\phi \\ \cos\delta \end{pmatrix} \tag{2.6}$$

$$\hat{\mathbf{d}} = \begin{pmatrix} \sin\lambda\cos\delta\sin\phi + \cos\lambda\cos\phi \\ \sin\lambda\cos\delta\cos\phi - \cos\lambda\sin\phi \\ \sin\lambda\sin\delta \end{pmatrix} \tag{2.7}$$

$$\hat{\mathbf{b}} = \begin{pmatrix} \cos\lambda\cos\delta\sin\phi - \sin\lambda\cos\phi \\ \cos\lambda\cos\delta\cos\phi + \sin\lambda\sin\phi \\ \cos\lambda\sin\delta \end{pmatrix} \tag{2.8}$$

The slip vector $\hat{\mathbf{d}}$ is in the plane described by the normal vector $\hat{\mathbf{n}}$, thus both vectors are orthogonal and their dot product vanishes, i.e. $\hat{\mathbf{n}} \cdot \hat{\mathbf{d}} = 0$. It follows from the orthogonality of the nodal planes, that the normal vector of one nodal plane is the slip vector of the other and vice versa (e.g. *Stein and Wysession*, 2003). The two nodal planes of a FMS have subscripts 1 and 2:

$$\hat{\mathbf{n}}_1 = \hat{\mathbf{d}}_2 \quad \text{and} \quad \hat{\mathbf{d}}_1 = \hat{\mathbf{n}}_2 \tag{2.9}$$

Due to this symmetry, the nodal planes given by an FMS cannot be separated into rupture and auxiliary planes. And due to the exchangeability of both vectors follows for the null axes:

$$\begin{aligned} \hat{\mathbf{b}}_1 &= \hat{\mathbf{n}}_1 \times \hat{\mathbf{d}}_1 \\ &= \hat{\mathbf{d}}_2 \times \hat{\mathbf{n}}_2 = -\hat{\mathbf{b}}_2 \end{aligned} \tag{2.10}$$

Following the Wallace-Bott hypothesis the shear vector $\mathbf{s}$ is in the same direction as the fault slip $\hat{\mathbf{d}}$. This relation shows that the slip $\mathbf{d}$ can be considered as a function of the stress tensor and the fault surface normal which in turn is a function of strike and dip

$$\mathbf{d}(\phi, \delta, \lambda) = \mathbf{d}(\phi, \delta, \boldsymbol{\sigma}) = \mathbf{d}(\mathbf{n}, \boldsymbol{\sigma}) \tag{2.11}$$

The rake λ can be regarded as a function of strike, dip and stress orientations. Consider a given stress tensor where the orientation of an arbitrary intermediate stress $\hat{\mathbf{s}}_2 = \hat{\mathbf{b}}$ is fixed. Then the following equation can be derived from Eq. 2.8 (see

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

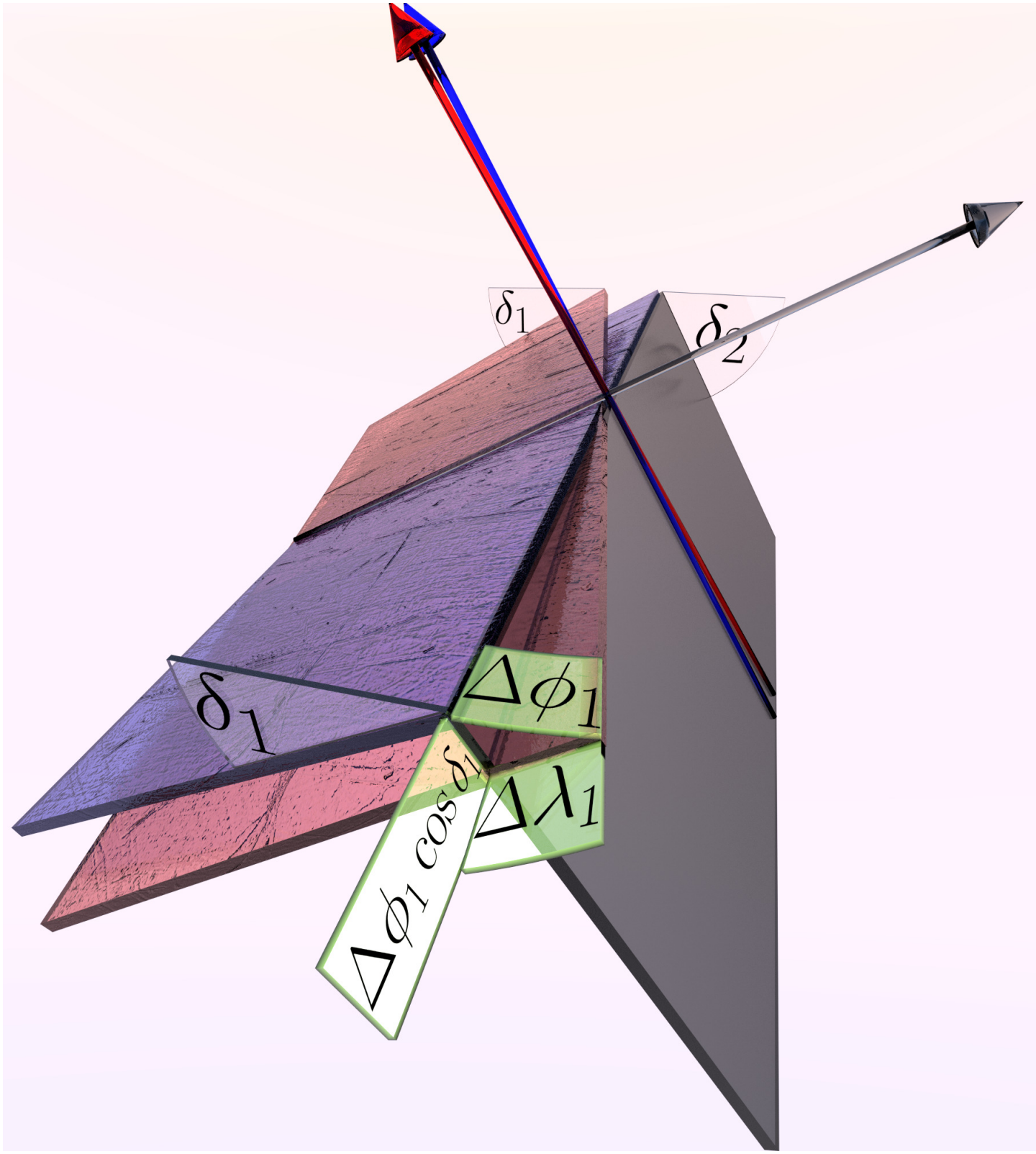


Figure 2.3: Model illustrating the relation between strike and rake of a nodal plane (red & blue rough surfaces, angles with index 1) for a constant slip vector (gray vector, normal of the smooth gray plane, index 2). If a nodal plane's strike ϕ_1 is shifted by an infinitesimal increment $\Delta\phi_1$, then the rake λ_1 is shifted by $\Delta\lambda_1$. The strike is defined on a horizontal plane, while the rake is defined on the nodal plane itself. The distance between the horizontal and the nodal plane shifted by $\Delta\phi_1$ and $\Delta\lambda_1$, respectively, is $\Delta\phi_1 \cos\delta_1$, i.e. $\Delta\lambda_1(\Delta\phi_1)^{-1} = \cos\delta_1$.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

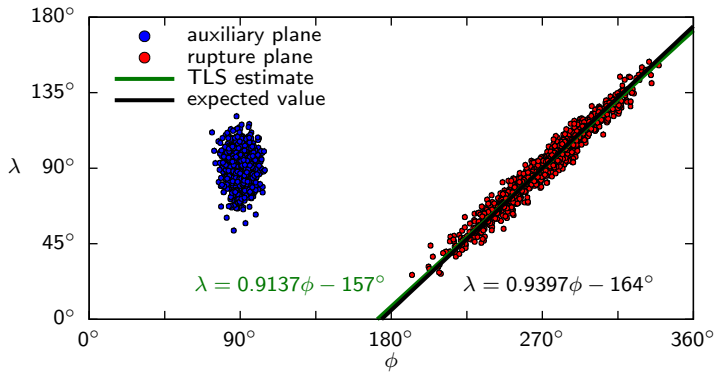


Figure 2.4: Shape dependence of nodal plane clusters. The clusters represent both nodal planes expressed as angles simulated from a near isotropic stress tensor. The variance is caused by adding normally distributed noise to the principal stress axis orientations. The simulated regime is reverse with maximum horizontal (and principal) stress orientation in north-south direction, resulting in an average strike and rake of 90°. The underlying fracture criterion is the Mohr-Coulomb criterion with a friction coefficient of 0.7. The first simulated rupture plane is reactivated in subsequent rupture events, therefore simulating the reactivation of faults. The blue cluster represents here the auxiliary plane and has a mean dip of 70°. Since the dip is close to zero, no relation between strike and rake is observable. The corresponding cluster of the rupture plane is the elongated cluster to the right (red). Due to its shallow mean dip of 20°, a strong relation between strike and rake is apparent. The slope of the cluster in the strike-rake plane, estimated by total least squares (TLS, green line), is close to a slope of $\cos 20° \approx 0.9397$ (black line), which is the expected value of the slope from Eq. 2.12.

appendix A.1.1 for derivation):

$$\frac{\partial \lambda(\phi, \delta)}{\partial \phi} = \cos \delta \tag{2.12}$$

$$\frac{\partial \lambda(\phi, \delta)}{\partial \delta} = 0 \tag{2.13}$$

The same relation holds for a constant slip vector $\hat{\mathbf{d}}$ of either nodal plane, then the derivatives in Eq. 2.12 and 2.13 are found by setting the derivative of Eq. 2.7 to zero (see appendix A.1.1). Figure 2.3 graphically visualizes the derivation of the derivatives. The two derivatives imply that in general on one side no functional relation exists between rake and dip (as is the case for strike and dip). On the other side a relationship exists between strike and rake according to Eq. 2.12. In Figure 2.4 we show FMS nodal plane clusters from synthetic data. Both relations between the angles can also be observed in global data sets, as shown for the entire GCMT catalog in Fig. 2.5.

The implications of the derivatives for the data distribution are taken into consideration for the cluster analysis. The cluster analysis itself is inductive, i.e. nothing is known about the clusters a priori. Any information about the clusters is derived during the run of the analysis. The advantage of not using a priori information comes with its adaptability to the data. Thus additional information, which possibly eludes a deterministic approach for investigating individual events, can be observed. This additional information from the clustering can be used to assign weights to FMS used in other applications, e.g. stress tensor inversion, b-value estimation, regression of ground-motion prediction equations etc.

2.2 Methods

In this section, we introduce the algorithm for identifying clusters of nodal planes from a FMS-catalog: **ACE** - **A**ngular **C**lassification with **E**xpectation-Maximization. As the acronym implies, the algorithm is based on the expectation-maximization (EM) algorithm (*Dempster et al.*, 1977).

We first introduce the statistical model to describe the data of a nodal plane population (2.2.1). This statistical model is a mixture model of probability distributions - with each component describing one cluster (in the ideal case, otherwise several components describe one cluster). The parameters of the mixture model are determined with the EM-algorithm. Since the cluster population size is unknown, we introduce two additional parameters to control the merging and removal of mixture components at each iteration of the EM-algorithm (section 2.2.2). Based on information theory, we present the optimization of these control parameters to determine the optimal number of components (section 2.2.3). Following in section 2.2.4, we identify subpopulations of nodal plane clusters from the mixture according to their Style-of-Faulting. The cluster subpopulations are the final outcome of ACE and are the basis for subsequent analyses. As an application example we show a weighted stress tensor inversion based on the works of (*Michael*, 1984; *Hardebeck and Michael*, 2006). The weights provided by ACE reduce the uncertainties of the inversion. We also investigate the Kagan angle distribution of the cluster subpopulations.

2.2.1 Statistical model of a nodal plane cluster

A focal mechanism is described as a set of three angles $\boldsymbol{\theta} = (\phi, \lambda, \delta)$, consisting of strike ($\phi$), rake ($\lambda$) and dip ($\delta$). Each focal mechanism is fully represented by two angle triples $\boldsymbol{\theta}$, describing the orientation of both nodal planes (i.e. rupture plane and auxiliary plane), respectively.

The last term in Eq. 2.2 is used to describe purely randomly distributed stresses. In order to cover this randomness we use a statistical model to describe the data distribution. The model for one cluster is denoted by the probability density function (PDF) $\mathcal{J}(\boldsymbol{\theta})$. This distribution has limited support in all three dimensions: $\phi \in [0, 360°)$, $\lambda \in [-180°, 180°)$, $\delta \in [0, 90°]$. Additionally, the angles show periodic behavior,

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

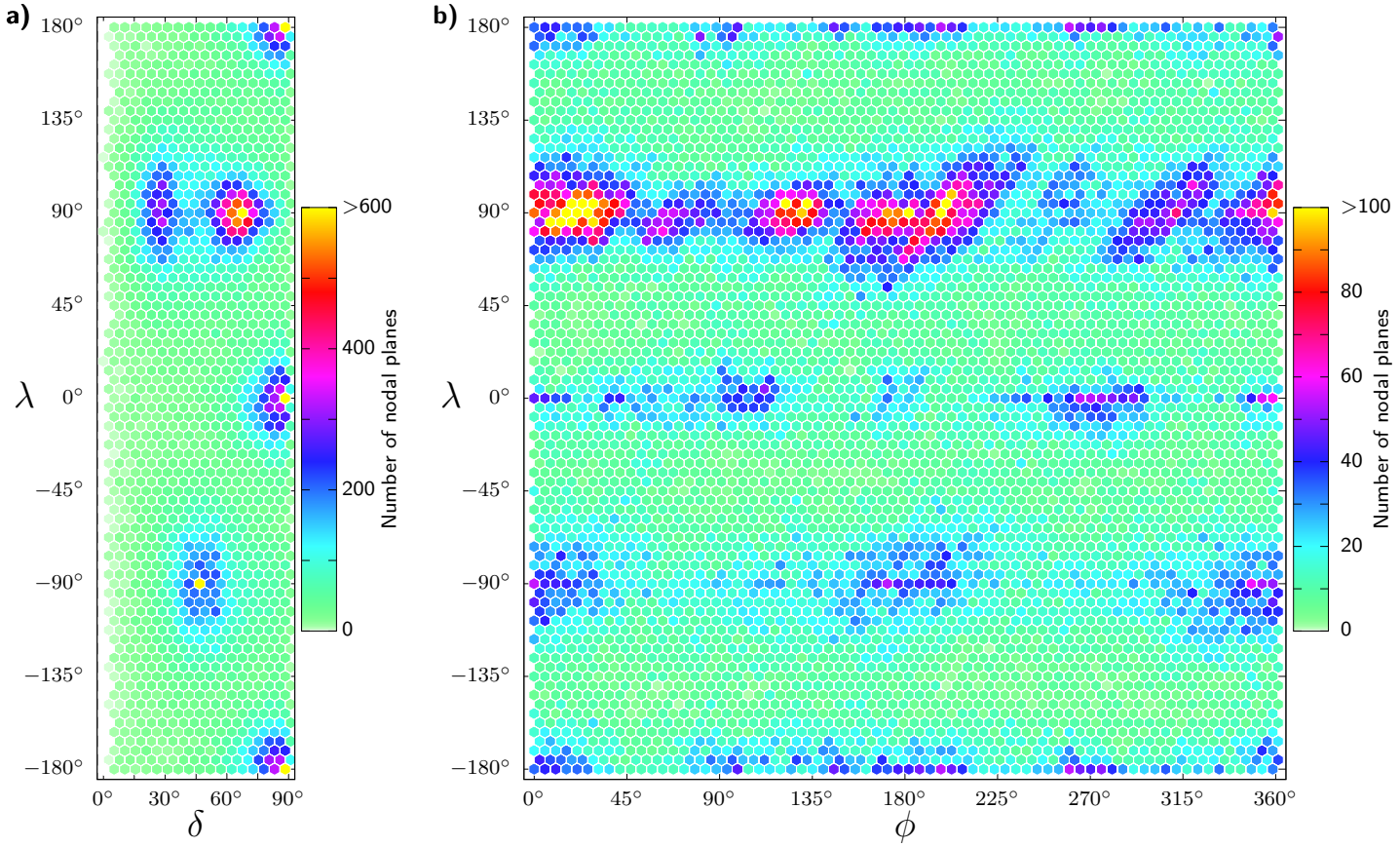


Figure 2.5: a) Distribution of dip (δ) and rake (λ) for the entire GCMT catalog from 1976 to 2016 for both nodal planes ($> 90,000$). Most nodal planes cluster at rakes of $\lambda = \{0°, \pm 90°, \pm 180°\}$, i.e. most events are close to Style-of-Faulting end members of normal, reverse and strike-slip. As indicated by Eq. 2.13 no apparent relation is visible in the distribution. b) Distribution of strike (ϕ) and rake (λ) for the same catalog as in a). Due to the predominant north-south striking of most continental collision zones and mid-ocean ridges, normal and reverse faulting ($\lambda = \pm 90°$) events cluster at strikes of $\phi = 0°$ and $\pm 180°$. Strike-slip events which commonly occur at transform faults perpendicular to the mid-ocean ridges have strikes of $\phi = \{0°, 90°, 180° 270°\}$. Several clusters of reverse faulting events ($\lambda = 90°$) show a linear behavior as indicated by Eq. 2.12 and shown in Fig. 2.4. This linear relationship becomes even more apparent when smaller regions are investigated (Fig. 4.3a, 4.7a, 4.10a).

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

which is best illustrated by the mean of the angles $(5°, 355°)$, which is not 180° but 0°. A distribution showing such behavior, i.e. a distribution around a circle (for the 1-dimensional case) is called a wrapped probability distribution (e.g. *Mardia and Jupp*, 1999).

Any PDF $P(x)$ can be converted to a wrapped PDF $P_w(\omega)$ by wrapping its variable and summing over the period 2π:

$$\omega = x \bmod 2\pi \qquad f_w(\omega) = \sum_{u=-\infty}^{\infty} f(\omega + 2\pi u) \tag{2.14}$$

The distribution of strike, rake and dip is three-dimensional and the model is placed in a model space defined on a 3-dimensional torus (also known as hypertorus; a four-dimensional doughnut, figuratively), which is the product of the three circles each defining the domain of strike, rake, and dip (though the dip spans only a quarter circle by definition). As already shown in the previous section, the rake and the strike of FMS are dependent on each other (Eq. 2.12) if the intermediate stress orientation (cross-product of normal and slip vector) or the slip vector is constant. We refer to a "constant stress orientation" for the first three components of Eq. 2.2 only, the random component is never considered constant, thus the vector orientations themselves are not completely constant.

The distribution of dip and rake of both nodal planes in the complete FMS catalog of the Global Centroid Moment Tensor (GCMT) catalog (*Dziewonski et al.*, 1981; *Ekström et al.*, 2012) from 1976 - 2016 shows, that earthquakes tend to cluster around $\lambda = \{0°, \pm 90°, \pm 180°\}$ (Fig. 2.5a). The distribution shows no correlation between rake and dip for all clusters of the Style-of-Faulting end members as is implied by Eq. 2.13. The independence of the dip from either strike and rake allows to define the distribution as a product:

$$\mathcal{J}(\phi, \lambda, \delta | \mu, \nu, \sigma^2, \tau^2, r, \alpha, \beta, a, b) = \mathcal{N}_w(\phi, \lambda | \mu, \nu, \sigma^2, \tau^2, r) \mathcal{B}(\delta | \alpha, \beta, a, b), \tag{2.15}$$

where $\mathcal{N}_w(\phi, \lambda)$ is the distribution in the strike-rake plane, and $\mathcal{B}(\delta)$ is the distribution of the dip. An advantage of this model is its separability into a wrapped and an unwrapped distribution. The distribution $\mathcal{N}_w(\phi, \lambda)$ is a bivariate wrapped normal distribution and cannot separated further into two independent distributions due to the dependence introduced by Eq. 2.12, which becomes most apparent for shallow dipping ruptures (e.g. thrust faulting, Fig. 2.5b). However, this distribution covers also those cases where strike and rake are independent from each other, as expected for strike-slip faulting with (nearly) vertically dipping nodal planes.

The dip is defined on the finite interval $[0°, 90°]$, i.e. the dip does not show the same periodicity as both the strike and rake. The distribution of the dip is modeled by a beta distribution $\mathcal{B}(\delta)$. The two distributions and their parameters are described in the following section.

2.2.1.1 Wrapped normal distribution

The bivariate wrapped normal distribution for the strike ϕ has mean μ and variance σ^2, the distribution for the rake λ has mean ν and variance τ^2 and the bivariate wrapped normal distribution can be written as

$$\mathcal{N}_w(\phi, \lambda | \mu, \nu, \sigma^2, \tau^2, r) = \sum_{u,v \in \mathbb{Z}} \mathcal{N}(\phi + 2\pi u, \lambda + 2\pi v | \mu, \nu, \sigma^2, \tau^2, r) \tag{2.16}$$

where $\mathcal{N}(\cdot)$ is a normal distribution (see Eq. A.17 in appendix A.2). According to Eq. 2.12, a correlation between ϕ and λ exists, which is expressed by the angular correlation coefficient r (*Fisher and Lee*, 1983).

The reasons to choose this distribution for describing the strike-rake distribution are the following:

1. All parameters in Eq. 2.16 have analytic representations of their estimators based on maximum likelihood (see appendix A.2 for a more detailed description). This contrasts with the widely used von Mises-Fisher distribution, whose parameters - specifically its concentration and in case of the more general Kent distribution the correlation as well) - can be estimated only iteratively, therefore increasing both inaccuracy and computational time.

2. It is a stable distribution, i.e. a linear combination of independent samples of that distribution results in the same distribution, though with different parameters. This means that e.g. when identifying a cluster with a single distribution even though that cluster originates from two independent (yet similar) distributions, the single distribution is nevertheless a reasonable description of the data. A similar argument for the stable distribution is given by *Kagan* (1990, 1992) to choose the Cauchy distribution (also a stable distribution) to describe the FMS rotation distribution.

3. Inherent errors of the FMS data are modeled by the closely related von Mises-Fisher distribution (*Silver and Jordan*, 1982). Even though we do not include errors of the FMS data directly, we assume that the data distribution is influenced by the von Mises-Fisher distribution.

From the definition of the nodal planes in Eq. 2.6, Eq. 2.7 and Eq. 2.9, it follows that the sign of the rake for both nodal planes of an earthquake is always the same. This constraint also limits the extent of the distribution of nodal planes with a particular Style-of-Faulting along the rake. In conjunction with the relations between strike and rake given in Eq. 2.12, the expected maximum spread of any distribution in Eq. 2.16 along its principal axes in terms of the standard deviation is less then 60°. Therefore, even the widest spreading cluster to be expected can be sufficiently approximated by a small number of wrappings in Eq. 2.16. It is sufficient to use a wrapping in the range of $\{u, v\} = \{-2, -1, 0, 1, 2\}$.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

2.2.1.2 Beta distribution

The dip is modeled by a generalized beta-distribution with arbitrary interval $[a, b]$ and is given by

$$\mathcal{B}(\delta|\alpha,\beta,a,b) = \frac{(\delta-a)^{\alpha-1}(b-\delta)^{\beta-1}}{(b-a)^{\alpha+\beta-1}B(\alpha,\beta)}, \tag{2.17}$$

where α and β are the shape parameters, $B(\cdot,\cdot)$ is the beta function and a and b are the interval limits. We follow the definition of *Feller* (1971) and exclude the limits from the range, i.e. $a < \delta < b$, for numerical stability, as $\mathcal{B} \to \infty$ at the interval limits for $\alpha < 1$ and $\beta < 1$ (See appendix A.3 for a more detailed description of this distribution).

The beta distribution is chosen for the following reasons:

1. like the wrapped normal distribution (Eq. 2.16), the parameters in Eq. 2.17 have analytic representations of their estimators. In this case the estimators are based on the method of moments (see appendix A.3).
2. When $\alpha \approx \beta$ and $\alpha > 2$ the beta distribution is similar to the normal distribution.
3. The distribution has a wide variety of shapes described by two parameters allowing data adaptabilty without resorting to distributions with large number of parameters.
4. Its shape variability allows also for fitting data with skewed distributions, e.g. dips close to 90° in case of strike-slip events.

2.2.1.3 Style-of-Faulting Index ($SoFI$)

Because of the same sign of the rake λ for both nodal planes for any FMS, the joint distribution $\mathcal{J}$ along the rake is truncated if the cluster consists mostly of either reverse or normal faulting nodal planes. Strike-slip events are an exception, as they cluster around $\lambda = 0°$ or $\lambda = \pm 180°$ and thus contain FMS with rakes of both signs (though both rakes of individual FMS have the same sign). In order to assess the SoF in which $\mathcal{J}$ is located, we introduce the Style-of-Faulting-Index $SoFI$ based on Eq. 2.6, 2.7, 2.8

$$SoFI = \frac{\hat{\mathbf{e}}_3^T \mathbf{A} \hat{\mathbf{e}}_3}{\hat{\mathbf{e}}_3^T \mathbf{C} \hat{\mathbf{e}}_3}, \tag{2.18}$$

where

$$\mathbf{A} = \hat{\mathbf{n}}\hat{\mathbf{d}}^T + \hat{\mathbf{d}}\hat{\mathbf{n}}^T \tag{2.19}$$

$$\mathbf{C} = \hat{\mathbf{n}}\hat{\mathbf{n}}^T + \hat{\mathbf{d}}\hat{\mathbf{d}}^T + \hat{\mathbf{b}}\hat{\mathbf{b}}^T \tag{2.20}$$

$$\hat{\mathbf{e}}_3 = (0,0,1)^T, \tag{2.21}$$

The definition of $SoFI$ is identical for both nodal planes, i.e. $SoFI$ is event based rather than nodal plane based.

Plugging the angle based definitions of the normal and slip vectors from Eq. 2.6 and 2.7 in Eq. 2.18 simplifies to

$$SoFI = \sin\lambda \sin 2\delta \tag{2.22}$$

According to Eq. 2.22, $SoFI$ is positive in a reverse tectonic regime, negative in a normal tectonic regime and around zero in a strike-slip tectonic regime. The $SoFI$ for the three end members of the tectonic regimes are

$$SoFI = \begin{cases} \frac{1}{2}\sqrt{3} & \text{if } \lambda = 90°, \delta = 30° \text{ (reverse)} \\ -\frac{1}{2}\sqrt{3} & \text{if } \lambda = -90°, \delta = 60° \text{ (normal)} \\ 0 & \text{if } \lambda = \{0°, \pm 180°\}, \delta = 90° \text{ (strike-slip)} \end{cases} \tag{2.23}$$

The regime index of distribution $\mathcal{J}$ indicates the predominant regime of that distribution and is calculated as a weighted mean of the individual $SoFI$ of the $2N$ nodal planes (N - number of events) with weights based on $\mathcal{J}$.

$$SoFI = \frac{\sum_{i=1}^{2N} \mathcal{J}_i SoFI_i}{\sum_{i=1}^{2N} \mathcal{J}_i} \tag{2.24}$$

We use this index as the indicator whether to apply truncation along the rake. The truncation of a distribution is applied, if $|SoFI| > 0.25\sqrt{3}$. The threshold is half the absolute value of $SoFI$ for the end members of reverse or normal faulting. In order to compensate for the truncation of $\mathcal{J}$ and to preserve the second axiom of probability - here specifically in the form of

$$\int \mathcal{J} d\boldsymbol{\theta} = 1 \tag{2.25}$$

the wrapped normal distribution with truncation along λ is

$$\mathcal{N}_{w,T}(\phi,\lambda|\mu,\nu,\sigma^2,\tau^2,r) = \begin{cases} \frac{1}{F_T} \sum_{u,v\in\mathbb{Z}} \mathcal{N}(\phi+2\pi u, \lambda+2\pi v|\mu,\nu,\sigma^2,\tau^2,r) & \text{if } |SoFI| > \frac{\sqrt{3}}{4} \\ 0 & \text{if } |SoFI| \le \frac{\sqrt{3}}{4} \end{cases} \tag{2.26}$$

where $1/F_T$ is the truncation factor with

$$F_T = \frac{1}{2} \sum_{v\in\mathbb{Z}} \operatorname{erf}\left(\frac{\pi + 2v\pi - |\nu|}{\tau\sqrt{2}}\right) - \operatorname{erf}\left(\frac{2v\pi - |\nu|}{\tau\sqrt{2}}\right) \tag{2.27}$$

The $SoFI$ from Eq. 2.22 in Eq. 2.26 for a distribution is estimated by

$$SoFI = \sin\nu \sin\frac{\pi\alpha}{\alpha+\beta}, \tag{2.28}$$

i.e. the means of λ and δ. This definition corresponds to the expression in Eq. 2.24 for the average .

The statistical model in Eq. 2.29 using the truncated wrapped normal distribution in Eq. 2.26 is stated as

$$\mathcal{J}_T(\phi,\lambda,\delta|\mu,\nu,\sigma^2,\tau^2,r,\alpha,\beta,a,b) = \mathcal{N}_{w,T}(\phi,\lambda|\mu,\nu,\sigma^2,\tau^2,r)\mathcal{B}(\delta|\alpha,\beta,a,b), \tag{2.29}$$

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

and is the basis for the mixture model and the EM algorithm in the next subsection. For brevity and readability, the subscript T is dropped in the following and the usage of the truncated distribution is implied in all instances.

2.2.2 Identification of nodal plane clusters with the EM algorithm

The EM algorithm iteratively computes the maximum-likelihood estimates in the presence of incomplete data, i.e. some latent variables which are only indirectly detectable in the observations (*Dempster et al.*, 1977). Each iteration of the EM-algorithm consists of two steps: (1) Expectation-step - Estimation of the expected value of the likelihood for the current estimate of the parameters of the statistical model, given the observations. (2) Maximization-step - Calculation of parameters maximizing the likelihood. The iteration process is repeated until the results converge.

The joint distribution $\mathcal{J}$ describes one component of a mixture model, and ideally represents one cluster within the data. Since a population of clusters is present in the data, the parameters of the components that describe the clusters are the latent variables. Using EM, the parameters of the components can be estimated, by simultaneously maximizing their likelihood iteratively.

The kth component in a mixture of K components is described by $\mathcal{J}_k$ which is weighted by w_k. For all K weights holds

$$\sum_{k=1}^{K} w_k = 1, \tag{2.30}$$

and the data are described by the full mixture model

$$M_{\mathcal{J}} = \sum_{k=1}^{K} w_k \mathcal{J}_k. \tag{2.31}$$

Two typical problems of EM are (1) its inability to handle noise, i.e. scattered data not belonging to any cluster, and (2) the number of components must be known a priori.

For our purpose, the noise data problem is solved by designating a PDF specifically for the noise in the data. We assume "noisy data" to have a wrapped uniform distribution, which is a special case of both distributions $\mathcal{N}_w$ $(\sigma, \tau \to \infty)$ and $\mathcal{B}$ $(\alpha = \beta = 1)$ and thus of $\mathcal{J}$ (Eq. 2.32). A wrapped uniform distribution implies that events are caused by random tectonic forces, i.e. not only the random stress term in Eq. 2.2 is variable but the remaining three as well, most notably the tectonic and earthquake related stress terms. Therefore this component of the mixture model contrasts with the data represented by the other components, where the random stress term is the only variable one. We investigate the properties of the identified clusters in terms of the Kagan angle (section 3.2 and 4).

We handle the second problem of EM, the a priori knowledge of the component number, by running EM with parameters controlling the component number at each iteration. To summarize the purpose of these control parameters:

- Check for overlap between components and merge components, if a threshold is reached ($\to p$).
- Check for the size of componets and remove those falling below a threshold ($\to q$).

With the control parameters, the extended EM algorithm consists of three steps:

1. The expectation step is performed by calculating Eq. 2.31.
2. The calculated components are checked for their overlap and weight and finally
3. in the maximization step the parameters of the components are updated.

The algorithm is repeated until the model converges or a maximum number of iterations is reached (e.g. 100 is sufficient since in most cases the model converges relatively fast).

At initialization, EM requires parameters set a priori. For N events, there are $2N$ nodal planes and considering a special component for noise, $K = 2N + 1$ components in total are initialized, i.e. we assume each nodal plane as a component a priori. It is sufficient to use a representative subset for larger data sets with $> 2,500$ events, as the information gain per event decreases with each additional event in the catalog. The initial values of the kth component $\mathcal{J}_k$ are either based on basic assumptions or are estimated from the underlying data and thus reduce user interference to a minimum.

1. The strike and rake of the kth nodal plane angle triple are used as the means in the strike-rake plane, $\mu_k = \phi_k$ and $\nu_k = \lambda_k$).
2. Both variances of strike and rake are set to the same value based on a distance estimate from the minimum spanning tree of the data. See section 2.2.2.1. The variance should sufficiently cover a wider neighborhood around the center of each initial component.
3. No correlation between strike and rake is assumed, $r_k = 0$.
4. The initial component along the dip is modeled using a beta distribution kernel after *Chen* (1999). The mean of the kernel is δ_k. The variance (or bandwidth in terms of (*Chen*, 1999)), as for the strike and rake, is based on the distance estimate of the minimum spanning tree. See section 2.2.2.1
5. The SoF index is based on the rake and dip, $SoFI_k = \sin\lambda_k \sin 2\delta_k$

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

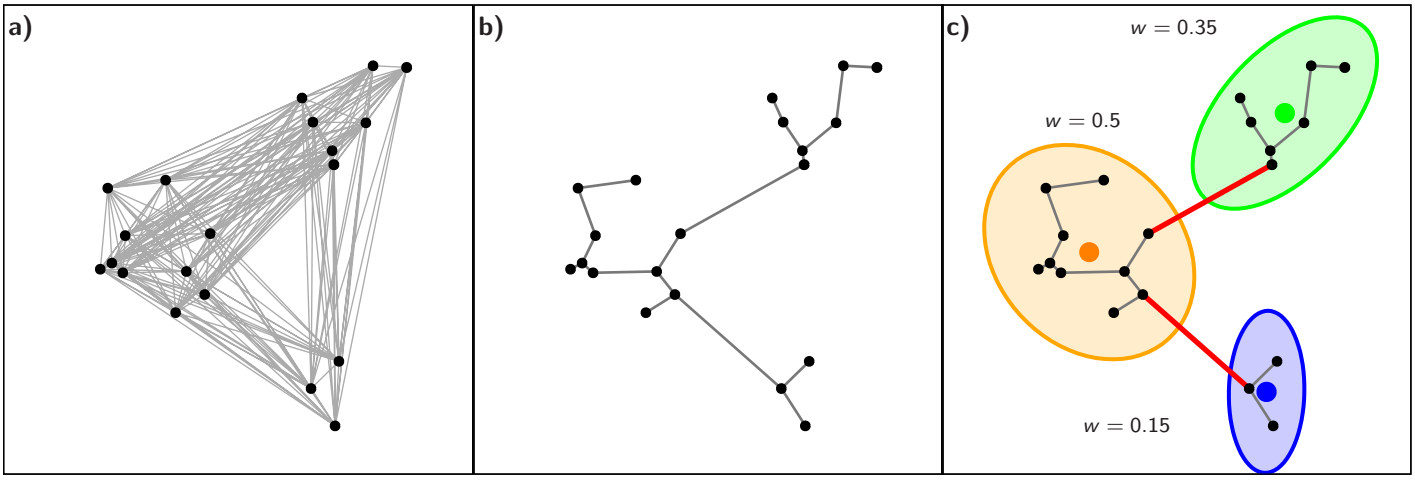


Figure 2.6: Sketch of the application of component merging and removal. a) In the first iteration of EM, all components have identical shape and their centroid is at the nodal plane coordinate (components shown by centroids only). The merging of two components requires a distance metric (gray lines) as criterion, which is the Jensen-Shannon distance. However, in order to merge two components according to control parameter p, not all Jensen-Shannon distances between all pairs are required. It is sufficient to use only a set of those component pairs which connect all component (directly and indirectly). This subset of pairs is found by Kruskal's algorithm (b) and represents those pairs connceting all componets with the shortest distance (minimum spanning tree). Next (c), all components with distances below p are merged into one new component (colored ellipses, red lines show distances larger than p). After the component merging, the weights, standing for the size of a component, are checked ($w < qK^{-1}$). The number of merged components here is $K = 3$, and with e.g. $q = 0.5$, the blue component is removed from the mixture (because $0.15 < 0.5/3$). In the subsequent iterations, the step involving Kruskal's algorithm becomes redundant and can be skipped (see text for explanation). This implies for subsequent iterations that only steps shown an (a) and (c) are performed.

6. The weight is the mean weight, $w_k = K^{-1}$.

The cluster designated for the noise (with index $k = K$) is initialized as a uniform distribution.

$$\begin{aligned}\mathcal{J}_K &= \mathcal{N}_w(\phi, \lambda|0, 0, \infty, \infty, 0)\mathcal{B}(\delta|1, 1, -1, 91) \\ &= \frac{1}{4\pi^2}\frac{23\pi}{45}\end{aligned} \tag{2.32}$$

Like the other distributions, the weight is at initiation the average weight $w_K = K^{-1}$.

Due to the large number of components at initialization ($K = 2N + 1$), it is very likely that neighboring components have a large overlap. The overlap - or similarity - of two components, represented by their PDF A and B with weights w_A and w_B, respectively, is measured by the Jensen-Shannon distance (*Österreicher and Vajda*, 2003)

$$\begin{aligned}d_{JS} = \Bigg[&- \frac{1}{w_A + w_B}\int w_A A \log_2 A + w_B B \log_2 B d\boldsymbol{\theta} \\ &- \int \left(\frac{w_A A + w_B B}{w_A + w_B}\right) \log_2 \left(\frac{w_A A + w_B B}{w_A + w_B}\right) d\boldsymbol{\theta}\Bigg]^{\frac{1}{2}}.\end{aligned} \tag{2.33}$$

The Jensen-Shannon distance is in the range between zero (identity of components) and one (completely dissimilar).

When two or more components have an overlap larger then a given threshold, they are merged into one component. The threshold of the Jensen-Shannon distance at which two components are merged is set by the control parameter p. Since p is not known, EM is performed for a set of thresholds p within range $0 < p < 1$.

Some components show little overlap with other components, but contain few events or only one. These remote minor components are undesired as they keep the total number of identified clusters high while containing little information. Therefore, components with weights below qK_i^{-1} are removed from the mixture. The value K_i^{-1} is the average component weight in the ith iteration. The weights are determined after components have been merged (Fig. 2.6c) in each iteration. The control parameter q is in the range $0 \leq q \leq 1$. With $q = 0$, no constraint is applied on the weights and with $q = 1$ total weight uniformity is implied. As with p, q is not known a priori and EM is run for a set of different values.

The overlap check and the weight check are at no instance applied to the designated noise component. It can thus be neither removed from the mixture model nor merged into another component. Consequently, no component can be merged into the noise component. If this special component is removed, another component (or several) following the model in Eq. 2.29 will cover the noisy data, which may in turn lead to overfitting. The only parameter controlling the noise

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

component is its weight, which is updated together with the weights of the other components at each iteration.

2.2.2.1 Application of the minimum spanning tree

The number of distances calculated in the first iteration is $N(2N-1)$. Because the Jensen-Shannon distance has no analytic expression, Eq. 2.33 is evaluated numerically. However, during the first iteration most distances are not required to merge components (Fig. 2.6a). Since all components share the same variances and weights during the first iteration, the Jensen-Shannon distance correlates with the angular euclidean distance between the means of the components (Fig. 2.7). The computationally much simpler euclidean distance is then used in Kruskal's algorithm (*Kruskal*, 1956) to identify those $2N-1$ component pairs for which the Jensen-Shannon distances are then calculated (Fig. 2.6b).

This reduction of calculations of the Jensen-Shannon distance is based on the triangle inequality:

$$d_{JS}(A,C) \leq d_{JS}(A,B) + d_{JS}(B,C). \tag{2.34}$$

If components A and C are not directly connected, then they are connected via component B (or for the general case, any number of components). Kruskal's algorithm ensures that all components are connected by the minimum sum of all distances, where for each component pair only one distance is given.

The (squared) angular Euclidean distance between two nodal planes is given by

$$\begin{aligned} d_{k,l}^2 = \quad & \left(e^{i\phi_k} - e^{i\phi_l}\right)^2 \\ & + \left(e^{i\lambda_k} - e^{i\lambda_l}\right)^2 \\ & + \left(e^{i\delta_k} - e^{i\delta_l}\right)^2 \end{aligned} \tag{2.35}$$

Given the extension of the nodal plane space, it holds

$$0 \leq d_{k,l} \leq \frac{3}{2}\pi \tag{2.36}$$

In subsequent iterations it is not possible to reduce the number calculations of the Jensen-Shannon distance, because both weights and variances vary and the correlation between the Jensen-Shannon and euclidean distance is not guaranteed anymore. Therefore the minimum spanning tree cannot be based on the Euclidean distance and must be based on the Jensen-Shannon distance, making the minimum spanning tree redundant because no computional advantage is achieved and all possible Jensen-Shannon distances are required due to the uniqueness of all componenets after the first iteration.

As stated in the list of initial condition, the distances given by minimum spanning tree provide an estimate of the initial variance for both the wrapped normal distribution and the beta distribution. The initial value is derived from the 95 % quantile of the empirical cumulative distribution function of the distances given by the minimum spanning tree.

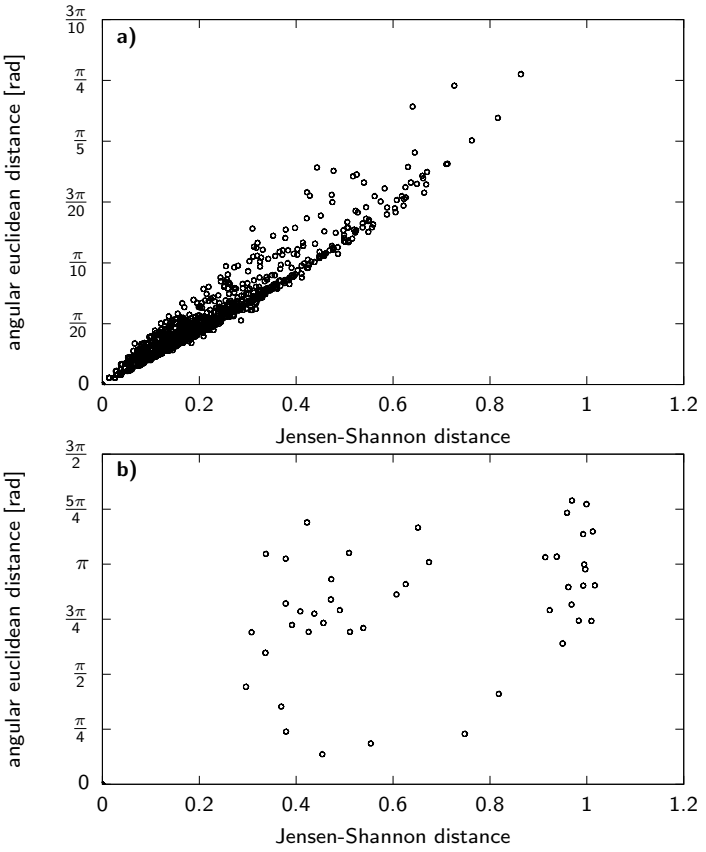


Figure 2.7: Jensen-Shannon distance (Eq. 2.33) versus angular euclidean distance (Eq. 2.35). The Jensen-Shannon distance is based on the pairs of the component probability densities, while the angular euclidean distance is based on the separation means of those components. a) Due to the uniform bandwidth (variance) and weights resulting in similarly shaped components during the first iteration the Jensen-Shannon distance correlates highly with the angular euclidean distance. This fact can be used to use the computionally much simpler angular euclidean distance and apply it with Kruskal's algorithm to significantly reduce the final number of Jensen-Shannon distances during the first iteration. In subsequent iterations (b, second iteration) the correlation between the two metrics disappears due to the adaption of the components to the data resulting in heterogenuously shaped components.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

2.2.3 Optimal parameters for controlling distribution merging and removal

Once the mixture models for different p and q have been calculated, an analysis based on information theory (*Shannon*, 1948) is applied to select the mixture model that describes the clusters of nodal planes. In the previous section we introduced the Jensen-Shannon distance, which is the square root of the Jensen-Shannon divergence for two distributions. The general form of the Jensen-Shannon divergence is (*Lin*, 1991)

$$D_{JS} = H\left(\sum_k w_k P_k\right) - \sum_k w_k H(P_k), \tag{2.37}$$

where w_k are weights with $\sum_k w_k = 1$. The probability distribution P_k has information entropy $H(P_k)$ (where H is the greek capital letter "eta"), which in case of a continuous distribution P is defined as (*Shannon*, 1948)

$$H(P) = -\int P(x) \log_b P(x) dx \tag{2.38}$$

specifically for the distributions of FMS angles (Eq. 2.29)

$$H(\mathcal{J}) = -\int \mathcal{J}(\boldsymbol{\theta}) \log_b \mathcal{J}(\boldsymbol{\theta}) d\boldsymbol{\theta}, \tag{2.39}$$

where the base b of the logarithm determines the range of the Jensen-Shannon divergence. An illustrating example for the Jensen-Shannon divergence is shown in Fig. 2.8.

If the number of distributions in Eq. 2.37 is equal to base b of the entropy, then the range of the Jensen-Shannon divergence is between 0 and 1. As for the Jensen-Shannon distance in Eq. 2.33, if $D_{JS} = 0$, then all distributions are identical, and if $D_{JS} = 1$ all distributions are fully dissimilar.

The Jensen-Shannon divergence is used to rank the mixture models according to the following criteria:

1. Orthogonality: If the slip vector of one nodal plane is the normal vector of the other plane (and vice versa), then the components representing these planes should behave in the same way (i.e. consistency with Eq. 2.9).

2. Goodness of fit: The mixture model should fit the observed data well.

3. Distinctiveness: Within a mixture model, the individual distributions should be as distinct as possible from each other. In other words, the components should be well separated, show little overlap, and easily distinguishable.

The basis for the first criterion is a change of variable. Here, the variables of a component of mixture model $M_{\mathcal{J}}$ are substituted by the variables of the other nodal plane. The mixture model with substituted variables is labeled $M^S_{\mathcal{J}}$. The change of variables for a probability density function is

$$P(\phi_2, \lambda_2, \delta_2) = P(\phi_1, \lambda_1, \delta_1) |\det \mathbf{J}| \tag{2.40}$$

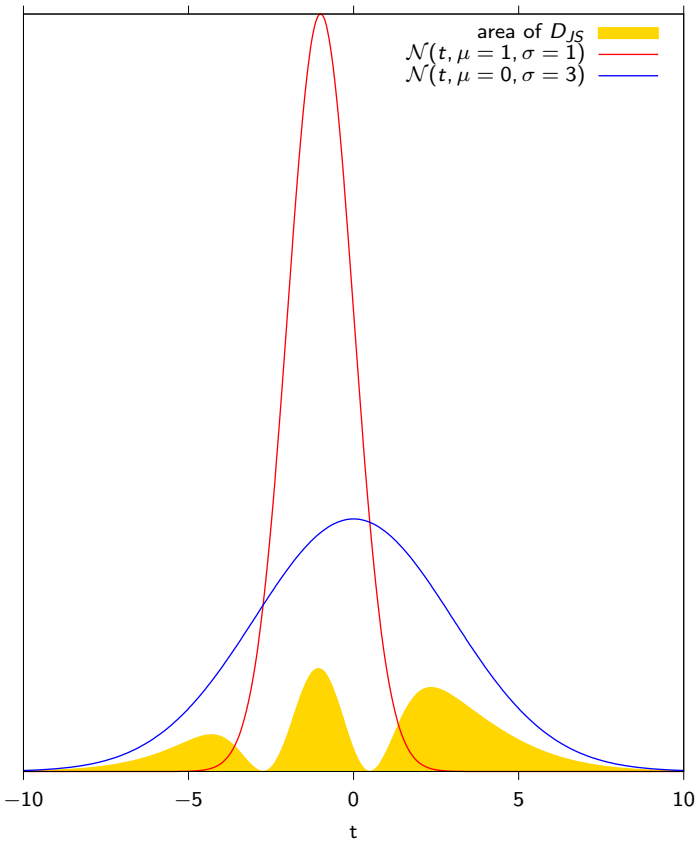


Figure 2.8: The Jensen-Shannon divergence measures the similarity between two distributions. The figure shows two normal distributions (blue, green). The Jensen-Shannon divergence corresponds to the yellow area. Note, how the yellow graph reaches zero where the two distributions intersect, while it increases where the two distributions differ more. It reaches its maximum where the two distributions are most different.

Assuming that the variables are changed from one nodal plane representation to the other nodal plane representation, the following relation holds (see appendix A.4 for derivation):

$$\mathcal{J}^S(\phi_2, \lambda_2, \delta_2) = \mathcal{J}(\phi_1, \lambda_1, \delta_1) \frac{\sin \delta_2}{\sin \delta_1} \tag{2.41}$$

The definition can also be expressed the other way around by exchanging the subscripts.

For the second criterion, it is necessary to represent the data in terms of probabilities of a non-parametric distribution. We use a kernel density estimator (KDE) (*Rosenblatt*, 1956; *Parzen*, 1962) similar to the distribution $\mathcal{J}$, with a kernel based on a wrapped bivariate normal distribution and a kernel for the beta distribution (*Chen*, 1999)with a bandwidth based on a constant variance for strike and rake, and a beta distribution specific bandwidth for the dip (see appendix A.5). The wrapping range is limited to $\{-1, 0, 1\}$ due to the selected small variance σ^2_{KDE} with respect to the whole

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

data range. The choice of the bandwidth is a common problem encountered with kernel density estimates (e.g. *Park and Marron*, 1990). If a constant variance is used in all entropy estimates, the KDE entropies will correlate with the variance, thus keeping the entropies proportional to each other for different variances.

We compare the mixture model with changed variables to the KDE as well, i.e. joining the first and second criterion. This additional constraint allows for the calculation of the divergence in one step. From Eq. 2.37, the divergence of the mixture model $M_{\mathcal{J}}$, the mixture model with changed variables $M^S_{\mathcal{J}}$, and the KDE h is measured by

$$D^M_{JS} = H\left(\frac{1}{3}\left(M_{\mathcal{J}} + M^S_{\mathcal{J}} + h\right)\right) - \frac{1}{3}\left[H(M_{\mathcal{J}}) + H(M^S_{\mathcal{J}}) + H(h)\right], \tag{2.42}$$

with

$$M_{\mathcal{J}} = \sum_{k=1}^{K} w_k \mathcal{N}_w(\phi_1, \lambda_1 | \mu_k, \nu_k, \sigma_k^2, \tau_k^2, r_k) \mathcal{B}(\delta_1 | \alpha_k, \beta_k) \tag{2.43}$$

$$M^S_{\mathcal{J}} = \sum_{k=1}^{K} w_k \mathcal{N}_w(\phi_2, \lambda_2 | \mu_k, \nu_k, \sigma_k^2, \tau_k^2, r_k) \mathcal{B}(\delta_2 | \alpha_k, \beta_k) \frac{\sin\delta_2}{\sin\delta_1} \tag{2.44}$$

$$h = \frac{1}{2N} \sum_{i=1}^{2N} N_w(\phi_1, \lambda_1 | \phi_i, \lambda_i, \sigma^2_{KDE}) \mathcal{B}(\delta_1 | \alpha_i, \beta_i) \tag{2.45}$$

where the logarithms of the entropies $H(\cdot)$ have base $b = 3$, and the divergence is in the range $[0, 1]$. According to the first two criteria, maximum similarity between the two mixture models and the KDE is desired, i.e. D^M_{JS} needs to be minimized.

For the third criterion, we compute the Jensen-Shannon divergence of all components of mixture model $M_{\mathcal{J}}$, D^C_{JS}. If all components are independent from each other, the Jensen-Shannon divergence reaches its maximum. Since the component describing the noise always overlaps with any other component, it would reduce D^C_{JS} and subsequently its resolving power. Therefore, the noise component (with index $k = K$) is excluded from the calculation of the divergence,

$$D^C_{JS} = H\left(\sum_{k=1}^{K-1} \frac{w_k}{1 - w_K} \mathcal{J}_k\right) - \sum_{k=1}^{K-1} \frac{w_k}{1 - w_K} H(\mathcal{J}_k), \tag{2.46}$$

where the logarithm of the entropies has base $b = K - 1$, and the divergence is in the range $[0, 1]$.

The mixture models associated with two divergences are ranked for all p and q by the following expression:

$$Q_{JS}(p, q) = D^M_{JS}(p, q)^{D^C_{JS}(p,q)} \tag{2.47}$$

The minimum rank jointly minimizes D^M_{JS} and maximizes D^C_{JS}, therefore the mixture model with smallest Q_{JS} best fulfills the three criteria outlined above simultaneously.

2.2.4 Subpopulations of nodal plane clusters with identical SoF

With the best mixture model $M^{opt}_{\mathcal{J}}$, we identify cluster subpopulations that form pairs of nodal planes according to Eq. 2.9. The subpopulations are identified by comparing each component to the component with changed variables (Eq. 2.41), as introduced in the previous section. The probabilities of clusters representing a nodal plane pair are very likely to correlate, while all other combinations are expected to be uncorrelated. The correlation between the probabilities of two distributions $\mathcal{J}_A(\boldsymbol{\theta})$ and $\mathcal{J}_B(\boldsymbol{\theta})$ is measured by the congruence coefficient (*Tucker*, 1951)

$$r_c = \frac{\sum_i^{2N} \mathcal{J}_A(\boldsymbol{\theta}_i) \mathcal{J}_B(\boldsymbol{\theta}_i)}{\sqrt{\sum_i^{2N} \mathcal{J}_A(\boldsymbol{\theta}_i)^2 \sum_i^{2N} \mathcal{J}_B(\boldsymbol{\theta}_i)^2}} \tag{2.48}$$

Since the probabilities are non-negative, the range of r_c is between zero and one.

The threshold for the congruence coefficients at which two clusters form a nodal plane pair is determined by a change point analysis. We apply a change point analysis similar to the autoregressive likelihood estimation for automated phase picking (*Pisarenko et al.*, 1987; *Kushnir et al.*, 1990). While the phase picking routine is based on a mixture of normal distributions, the distribution of the congruence coefficients is assumed to have arisen under a mixture of two beta distributions, with one distribution for uncorrelated and correlated congruence coefficients, respectively.

The expected number of components of the mixture model $M^{opt}_{\mathcal{J}}$ is low, and may be insufficient for a change point analysis to provide reliable results. Thus, we determine the congruence coefficients by bootstrapping (*Efron*, 1979) the component probabilities to obtain a sufficiently large sample of congruence coefficients. Bootstrapping not only increases the sample size for the change point analysis, but also includes the uncertainties of the congruence coefficients.

The change point analysis is based on Akaike's information criterion (AIC) (*Akaike*, 1974)

$$\mathrm{AIC} = 2k - 2\ln\mathcal{L} \tag{2.49}$$

where k is the number of free parameters and $\mathcal{L}$ is the likelihood of the statistical model, i.e. the mixture of beta distributions. The AIC is used to identify the best mixture of beta distributions. This is identical to identifying the preferred threshold, because the beta distributions are separated

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

at that threshold and otherwise independent from each other. The log-likelihood in Eq. 2.50 for a mixture of two beta distributions (Eq. 2.17) is

$$\ln \mathcal{L} = \sum_{i=1}^{T} \ln \mathcal{B}(r_{ci}|\alpha_l, \beta_l, 0, 1) + \sum_{i=T+1}^{N_r} \ln \mathcal{B}(r_{ci}|\alpha_u, \beta_u, 0, 1) \tag{2.51}$$

where N_r is the total number of congruence coefficients, T is the number of coefficients less than or equal the threshold, and α_l, β_l and α_u, β_u are the four shape parameters (therefore $k = 4$) estimated from the congruence coefficients below (first sum, subscripts with l) and above (second sum, subscripts with u) the threshold congruence coefficient, respectively. The AIC is determined for each congruence coefficient set as the threshold and the preferred threshold is identified where the AIC reaches a minimum.

All cluster combinations with congruence coefficients larger than the threshold are considered to form nodal plane pairs. Some clusters can be connected to more than one cluster, particularly in the case of strike-slip events, where one cluster will be generally connected to two others. This behavior for strike-slip events is caused by the additional strike ambiguity of subvertically dipping events. Up to four strike-slip clusters can form a subpopulation, while for reverse and normal faulting the size of the subpopulation is two.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Chapter 3

ACE & Stress

3.1 Application of ACE to stress mechanics

In the following section we show how to incorporate the results of ACE into applications. As examples, we provide the Kagan angle distribution and stress tensor inversion.

The normalized weight for the ith nodal plane angle triple of a subpopulation with L components from a mixture model with in total K components is given by

$$v_i = \frac{\sum_{l=1}^{L} w_l \mathcal{J}_l(\phi_i, \lambda_i, \delta_i)}{\sum_{k=1}^{K} w_k \mathcal{J}_k(\phi_i, \lambda_i, \delta_i)} \tag{3.1}$$

This normalized weight is the basis any further application.

3.2 Kagan angle distribution

Since ACE uses FMS, the final mixture model can be directly used to apply weights to the so-called Kagan angle. The Kagan angle is the smallest angle to rotate one FMS into another. It is determined by finding the rotation quaternion between two FMS that results in the smallest angle (*Kagan*, 2007). The distribution of this angle follows a rotational Cauchy distribution under the assumption of purely randomly distributed distortions (i.e. faults) in a constant stress field with random variations (Eq. 2.2).

In section 2.2.1, we introduced the definitions of the components of the mixture model. While we assume the joint probability distribution in Eq. 2.29 for all components, we use the special case of a uniform distribution of the angles to represent the FMS noise (i.e. those events not forming clusters) at all instances. In terms of the Kagan angle Φ, a uniform distribution of the angle triple (ϕ, λ, δ) implies a pure random distribution of double couple sources *Kagan* (1992). The equation of the rotational distribution for randomly distributed double couple orientations is given in *Kagan* (1992, Eq. (6) therein).

The rotational distributions have no general analytic representation, except for purely randomly distributed FMS (*Kagan*, 1992, 2007). To generate other rotational distributions we follow the procedure outlined by *Kagan* (1992). We focus here on the comparison between the rotations of the FMS data and some rotational distributions. Following *Kagan* (1992, 2013), we consider pure random DC (double couple) rotation distribution, the rotational Cauchy distribution and the rotational Fisher-Bingham distribution for comparison. The Fisher-Bingham distribution (*Kent*, 1982) is a generalization of the von Mises-Fisher distribution used in *Kagan* (1992, 2007) and is not limited to an isotropic distribution of normal/slip vectors, i.e. same variances for strike and dip. *Kent* (1982) defines the Fisher-Bingham distribution as

$$\mathcal{FB}(\mathbf{x}|\kappa, \beta, \mathbf{\Gamma}) \propto e^{\kappa \boldsymbol{\gamma}_1^T \cdot \mathbf{x} + \beta\left[(\boldsymbol{\gamma}_2^T \cdot \mathbf{x})^2 - (\boldsymbol{\gamma}_3^T \cdot \mathbf{x})^2\right]} \tag{3.2}$$
$$\text{with } 0 \leq \beta < \frac{\kappa}{2} \text{ and } \mathbf{\Gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3)$$

with $\mathbf{x}$ being the $\hat{\mathbf{n}}$, $\hat{\mathbf{d}}$, or $\hat{\mathbf{b}}$ axes; concentration parameter κ (analogue to an inverse variance), shape factor β and the orthogonal matrix $\mathbf{\Gamma}$ with vectors $\hat{\boldsymbol{\gamma}}_1$ (mean axis), $\hat{\boldsymbol{\gamma}}_2$ (major axis of shape ellipse) and $\hat{\boldsymbol{\gamma}}_3$ (minor axis of shape ellipse). For brevity we left the normalization constant of the distribution out, see *Kent* (1982) for a detailed description. For $\beta = 0$ the Fisher-Bingham distribution becomes the von Mises-Fisher distribution. Its rotational counterpart is denoted by $\mathcal{FB}_r(\Phi|\kappa, \beta)$ and for $\kappa = \beta = 0$ it becomes the random DC distribution.

To obtain samples for $\mathcal{FB}(\mathbf{x}|\kappa, \beta, \mathbf{\Gamma})$ we use acceptance-rejection sampling (*von Neumann*, 1951) by accepting a sample of the Fisher-Bingham distribution for the slip, normal and null vectors (Eq. 3.2) if all three are accepted simultaneously. The matrices $\mathbf{\Gamma}$ for each of the three vectors are defined such to represent the orthogonality of the three vectors.

The rotational distributions for the different FMS subpopulations are estimated by taking the probabilities of the mixture model for each angle triple in a subpopulation as weights. We denote the rotational angle between the FMS given by ith and jth angle triple as φ_{ij}. The weight for the rotation angle φ_{ij} is the product of the ith and jth weight from Eq. 3.1.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

With N angle triples and neglecting rotations between same angle triples, a total of $N(N-1)$ rotations exist. The empirical weighted cumulative distribution function in terms of the rotational angle Φ is

$$F(\Phi) = \frac{\sum_{i,j=1}^{N} v_i v_j \mathbf{I}(\varphi_{ij} \leq \Phi)}{\sum_{i,j=1}^{N} v_i v_j} \text{ with } i \neq j, \tag{3.3}$$

where $\mathbf{I}(x)$ is an indicator function that is one if condition x is true and zero otherwise. The range of the rotation angle is $0° \leq \Phi \leq 120°$ (*Kagan*, 1992).

3.3 Stress Tensor Inversion

Since ACE's theoretical basis is related to stress tensor inversion, the cluster results can be used as a priori information for a stress tensor inversion. The a priori information comes as a definition for FMS subpopulations. Within each subpopulation, e.g. for thrust faulting at a continental plate interface, the stress state can be assumed to be more consistent than for the overall data set. ACE's subpopulations can be seen as a way to treat a FMS catalog not as uniform data set for stress inversion, but as a polyphase data set (*Célérier et al.*, 2012, sec. 7.4 and references therein), i.e. a data set divided into consistent subsets, and stress is inverted from each subset to arrive at a more substantial picture of the stress state.

We use the least squares (LSQ) approach introduced by *Michael* (1984), nonetheless ACE's a priori information can be implemented in any other stress inversion technique. Micheal's method is defined by the following relation:

$$\mathbf{s} = \mathbf{A}\mathbf{t} \tag{3.4}$$

$\mathbf{s}$ is a vector containing N unit slip vectors and matrix $\mathbf{A}$ is based on N normal vectors and $\mathbf{t}$ is the reduced stress tensor in vector notation. Two fundamental assumptions are present in Eq. 3.4: The Wallace-Bott hypothesis holds and the slip magnitude is uniform in all directions and unknown (implied by using the unit slip vectors). The difference of the reduced (or deviatoric) stress tensor to the stress tensor in Eq. 2.1 is the absence of an isotropic stress component (pressure p). The deviatoric stress tensor $\boldsymbol{\sigma}^{dev}$ is defined by

$$\boldsymbol{\sigma}^{dev} = \boldsymbol{\sigma} - p\mathbf{I} \tag{3.5}$$

where $\mathbf{I}$ is the identity matrix and pressure p is

$$p = \frac{\mathrm{tr}(\boldsymbol{\sigma})}{3} = \frac{\sigma_{11} + \sigma_{22} + \sigma_{33}}{3}, \tag{3.6}$$

thus follows $\boldsymbol{\sigma}_{33}^{dev} = \boldsymbol{\sigma}_{11}^{dev} + \boldsymbol{\sigma}_{22}^{dev}$, i.e. the deviatoric stress tensor has five independent components only. For convenience the stress tensor is expressed in Voigt notation

$$\mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \end{pmatrix} = \begin{pmatrix} \sigma_{11}^{dev} \\ \sigma_{22}^{dev} \\ \sigma_{12}^{dev} \\ \sigma_{13}^{dev} \\ \sigma_{23}^{dev} \end{pmatrix} = \begin{pmatrix} \sigma_{11} - p \\ \sigma_{22} - p \\ \sigma_{12} \\ \sigma_{13} \\ \sigma_{23} \end{pmatrix} \tag{3.7}$$

The matrix $\mathbf{A}_i$ is based on the normal vector $\mathbf{n}_i$ (subscript i dropped in the following equation) and follows from Eq. 2.5

$$\mathbf{A} = \begin{pmatrix} n_1 - n_1^3 + n_1 n_3^2 & -n_2 n_1^2 + n_2 n_3^2 & -n_3 n_1^2 - n_3 + n_3^3 \\ n_2 - 2n_2 n_1^2 & n_1 - 2n_1 n_2^2 & -2n_1 n_2 n_3 \\ n_3 - 2n_3 n_1^2 & -2n_1 n_2 n_3 & n_1 - 2n_1 n_3^2 \\ -n_1 n_2^2 + n_1 n_3^2 & n_2 - n_2^3 + n_2 n_3^2 & -n_2^2 n_3 - n_3 + n_3^3 \\ -2n_1 n_2 n_3 & n_3 - 2n_3 n_2^2 & n_2 - 2n_2 n_3^3 \end{pmatrix}^T \tag{3.8}$$

Inverting for $\mathbf{t}$ results in

$$\mathbf{t} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{s} \tag{3.9}$$

This is the general solution of the overdetermined LSQ problem. Weighted LSQ introduces the weight matrix $\mathbf{W}$

$$\mathbf{t} = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\mathbf{s} \tag{3.10}$$

with weights based on Eq. 3.1

$$\mathbf{W} = \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_N \end{pmatrix} \tag{3.11}$$

If clusters are determined for larger regions it is reasonable to invert for a set of tensors that are linearly dependent in space (*Hardebeck and Michael*, 2006, modified from Eq. 14)

$$\mathbf{t}_a = (\mathbf{A}^T\mathbf{W}\mathbf{A} + \epsilon^2\mathbf{D}^T\mathbf{D})^{-1}\mathbf{A}^T\mathbf{W}\mathbf{s} \tag{3.12}$$

where $\mathbf{D}$ expresses the linear dependency of the neighboring tensors and ϵ is the dependency strength and $\mathbf{t}_a$ is the vector of all stress tensors.

$$\mathbf{t}_a = \begin{pmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_n \end{pmatrix} \tag{3.13}$$

Calculation of the stress tensors according to Eq. 3.10 and 3.12 are performed with the armadillo library for C++ (*Sanderson and Curtin*, 2016).

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Chapter 4

Examples

We present three examples of ACE for Northern Chile, the Nazca plate and Kyūshū (Japan). The examples show the best mixture models from ACE for the regions. A comparison of STI with and without ACE based weights are shown as well as the distribution of Kagan angles.

All results of ACE we present here are given as probabilities as defined in Eq. 3.1. In the figures showing the probabilities of the Style-of-Faulting, we use a color code similar to the World Stress Map, based on the HSV (hue, saturation, [brightness] value) color space. The hue represents the SoF, and brightness represents the probability of the given SoF (Fig. 4.1). While only one color hue is assigned for all events in a component, the brightness shows the probability of each nodal plane to be in that particular component. If several subpopulations of a similar (or even identical) SoF are present, colors are separated by saturation, i.e. one subpopulation appears generally lighter/darker than the other. When the hypocenter associated with a FMS is plotted, the color hue, saturation and lightness value are based on the average of both nodal planes. In case we cannot clearly differentiate between left- and right-lateral strike-slip, green is used.

4.1 Northern Chile

We apply ACE to focal mechanism data from the GCMT catalog (*Dziewonski et al.*, 1981; *Ekström et al.*, 2012) from 1976 - 2016 for northern Chile. This catalog contains 848 events, resulting in 1696 nodal planes. For northern Chile five clusters in three subpopulations have been identified, two clusters for normal and reverse faulting, respectively, and one cluster for unclassified SoF (Fig. 4.3a).

The reverse faulting clusters are associated with the plate interface of the colliding South American and Nazca plates, while the normal faulting clusters relate to deeper events in the downgoing slab (Fig. 4.3b,d). The unclassified events are mostly near the surface and are associated with thrusting in the Andes.

Several major events occurred in the region and their corresponding focal mechanisms (Fig. 4.3c) are in the centers of each identified cluster, respectively.

The distributions of the Kagan angle Φ for the Chilean FMS are shown in Fig. 4.2. The empirical cumulative probability distributions (ECDF) are calculated by Eq. 3.3 for each subpopulation in Fig. 4.3. Reverse faulting FMS (cyan curve) show high similarity between themselves as appr. 80 % of the FMS differ by a rotational angle of less than 30°. The normal faulting FMS (red curve) are more diverse and show larger rotational angles (nearly twice as much as for interface FMS), indicating a more complex stress field for these events which are also at larger depths compared to the reverse events. The remaining unclassified events (gray curve) are nearly completely randomly distributed, as their rotational distribution is similar to a purely random distribution of FMS (black curve). The three empirical distributions are fitted to rotational Fisher-Bingham distributions ($\mathcal{FB}_r$, dashed curves), parameters are given in Fig. 4.2. All empirical distributions are well described by the rotational Fisher-Bingham distribution.

Fig. 4.2 shows also the rotation angle distribution for unweighted data (light gray line). The empirical distribution is derived similarly to *Kagan* (1992). We consider all events in the region with a focal depth shallower than 100 km. Only rotations of FMS with a hypocentral distance of less than 50 km are taken into account. The resulting distribution indicates similarity for most shallow events in Northern Chile. This coincides with the empirical distribution of the interface subpopulation which is mostly at shallow depths (Fig. 4.3d). In agreement with *Kagan* (1992); *Kagan and Jackson* (2015) the unweighted distribution is well described a rotational Cauchy distribution (Fig. 4.2, dotted line).

Furthermore, the concentration parameter κ of the rotational Fisher-Bingham distribution correlates inversely with the variances of the wrapped normal distribution, i.e. the extent of the nodal plane clusters in Fig. 4.3a corresponds to the steepness of the ECDF in Fig. 4.2.

We conducted two stress tensor inversions, one, conventional inversion, using all events in the catalog of Northern Chile with an event depth of maximum 50 km. These shallow reverse faulting events are associated with interface activity. The second inversion uses the probabilities of the cluster subpopulations for reverse faulting as weights according to Eq.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

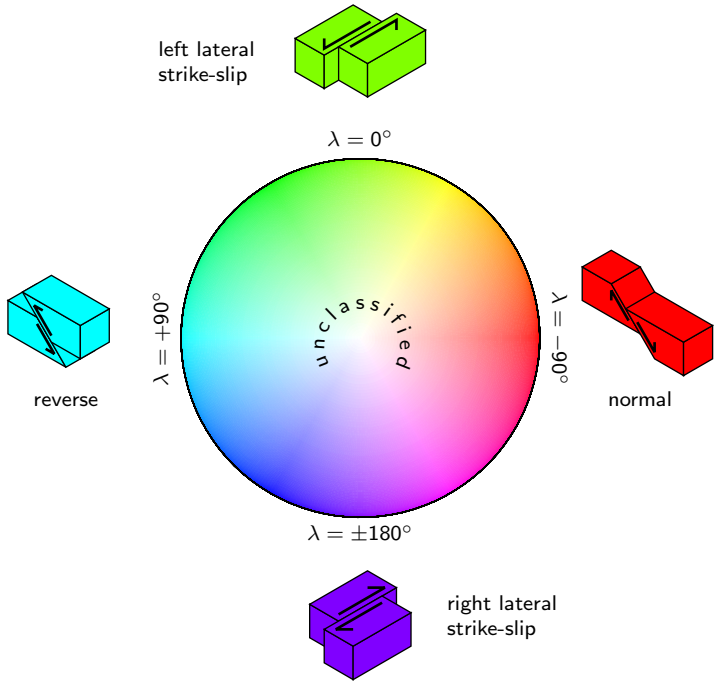


Figure 4.1: Color palette based on the rake to indicate the Style-of-Faulting of nodal plane clusters. The three SoF end members are also shown (for strike-slip both left- and right-lateral are shown). The rake hue corresponds to the hue of the HSL colorspace, with the end members assigned the colors chartreuse (left-lateral strike-slip), cyan (reverse), violet (right-lateral strike-slip) and red (normal). The color brightness corresponds to the probability of a nodal plane/event belonging to a certain cluster. The noise cluster has full brightness (i.e. white) and data in that range are unclassified. These color definitions are used in the subsequent figures.

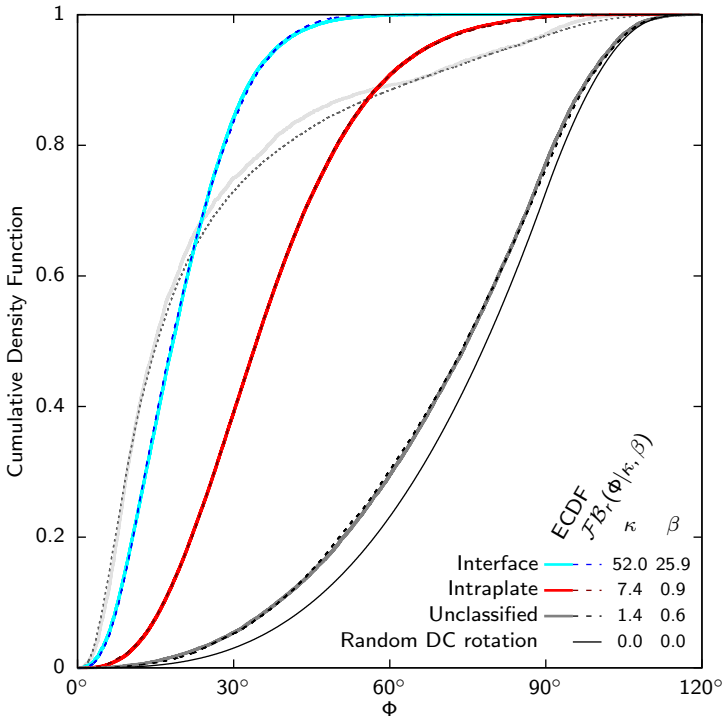


Figure 4.2: Cumulative distribution functions of the rotation angle between pairs of FMS from northern Chile. The colored solid curves are estimated from the FMS data by using the subpopulation probabilities as weights. The empirical distributions are well described by rotational Fisher-Bingham distributions ($\mathcal{FB}_r$, dashed lines, parameters are in table). The angle distribution for unclassified FMS is nearly random since it is close to the pure random DC distribution (thin black line). The solid light gray line shows the rotation angle distribution for pairs of FMS shallower than 100 km and with a hypocenters separated by less than 50 km. This unweighted distribution is best described by a rotational Cauchy distribution (κ_c=0.06, dotted line).

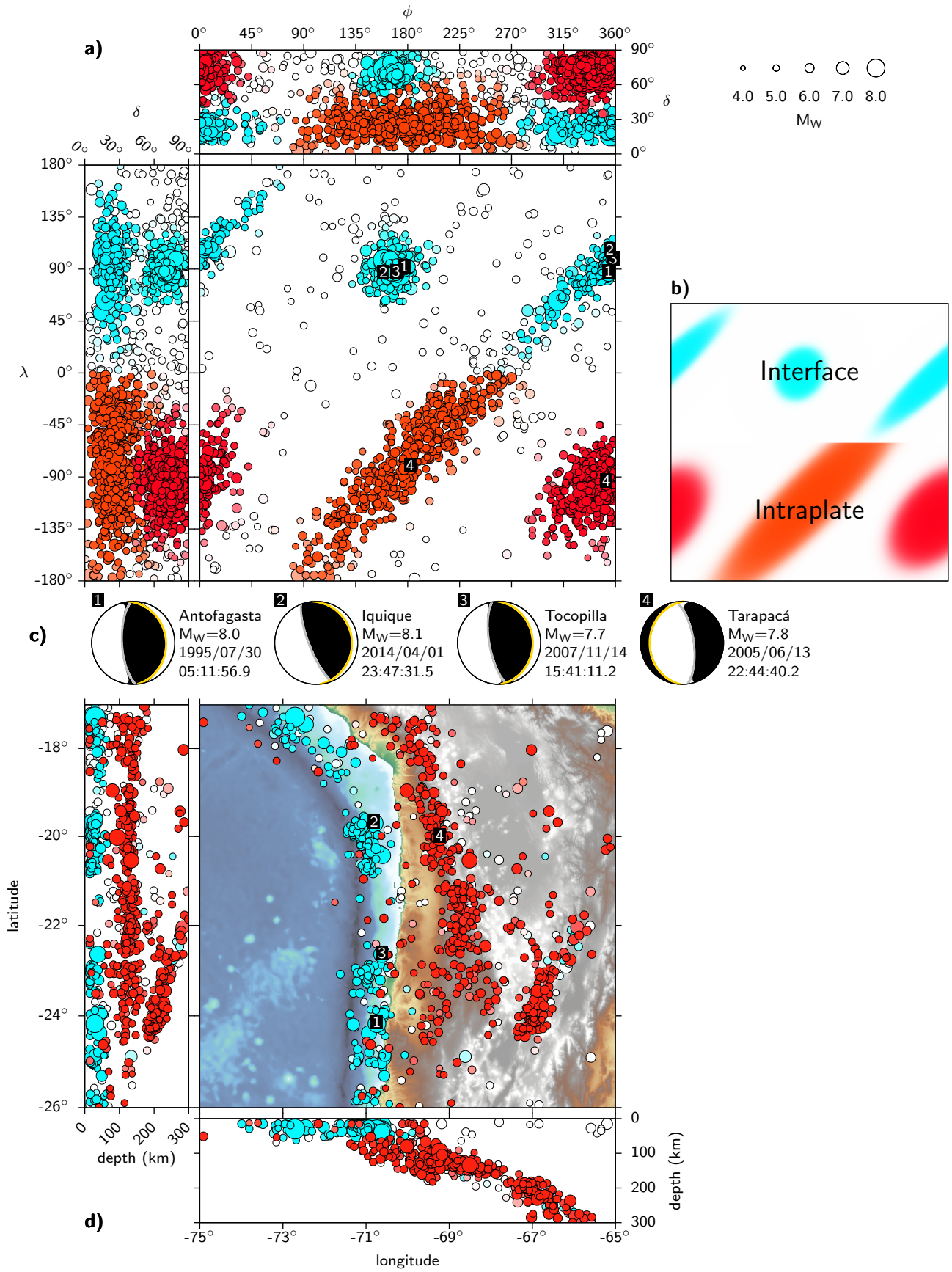
3.1. These events are also associated mostly with interface activity. Fig. 4.4a shows the results of the conventional stress tensor inversion, i.e. all events are equally weighted. The plot in Fig. 4.4 shows the distribution of 20,000 bootstrapped results using a kernel density estimator with a von Mises-Fisher distribution (concentration $\kappa = 2500$). The orientations of the principal stresses are shown in a lower hemispheric plot. While the minimum stress S_3 (cyan) is well located and nearly vertical, the horizontal stresses S_1 (magenta) and S_2 (yellow) are less well constrained. Uncertainties are estimated by bootstrapping following *Michael* (1987) by selecting randomly events and randomly setting one nodal plane is rupture plane and one as auxiliary plane.

The inversion results based on ACE are shown in 4.4b. While principal stress orientations are comparable to the conventional solution in Fig. 4.4, uncertainties are significantly reduced. Uncertainties are estimated with same procedure as for the conventional inversion.

Based on the motion direction of the Nazca plate, the clusters for interface events are separable into a cluster consisting of rupture planes ($\phi \approx 0°$,$\delta \approx 15°$) and auxiliary planes ($\phi \approx 180°$,$\delta \approx 75°$). For the normal faulting subpopulation no such disambiguation is feasible due to two-sided fault activation. The lack of preferred rupture orientation has been confirmed by aftershock analysis and directivity analysis (e.g. *Warren*, 2014).

Furthermore, the event hypocenters in the Northern Chile region identified as reverse faulting are in the upper 50 km, a common threshold for the occurrence of reverse faulting in that region (e.g. *Händel et al.*, 2014) and all large magnitude reverse events occurred in shallower depths (Fig. 4.3c).

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012



STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Figure 4.3: (left page) a) The optimal mixture model for the Northern Chile FMS data set. Two clusters for both normal and reverse faulting are identified. Both regimes show one (more or less) circular and one elongated cluster. b) Legend of the subpopulations. The reverse faulting clusters (cyan) are related to the interface of the colliding plates. The events in the normal faulting clusters (red) form the subpopulation associated with the subducting slab. FMS in the white area are unclassified. c) Four example beach ball plots of significant events in the region. The rupture plane and auxiliary planes are indicated as yellow and gray great-circles, respectively. d) Hypocentral locations of earthquakes from the GCMT catalog in Northern Chile. Nearly all reverse faulting events occur in the upper 50 km along the coast, while most normal faulting events are deeper in the slab. Few normal faulting events are close to surface along the shorelines. These events may be associated with the bending of the Nazca plate before being subducted. Many unclassified events are near the surface. These events are likely related to the thrusting of the Andes due to their locations, as well as their frequently positive rake values. Map data: ETOPO1 (*Amante and Eakins*, 2009).

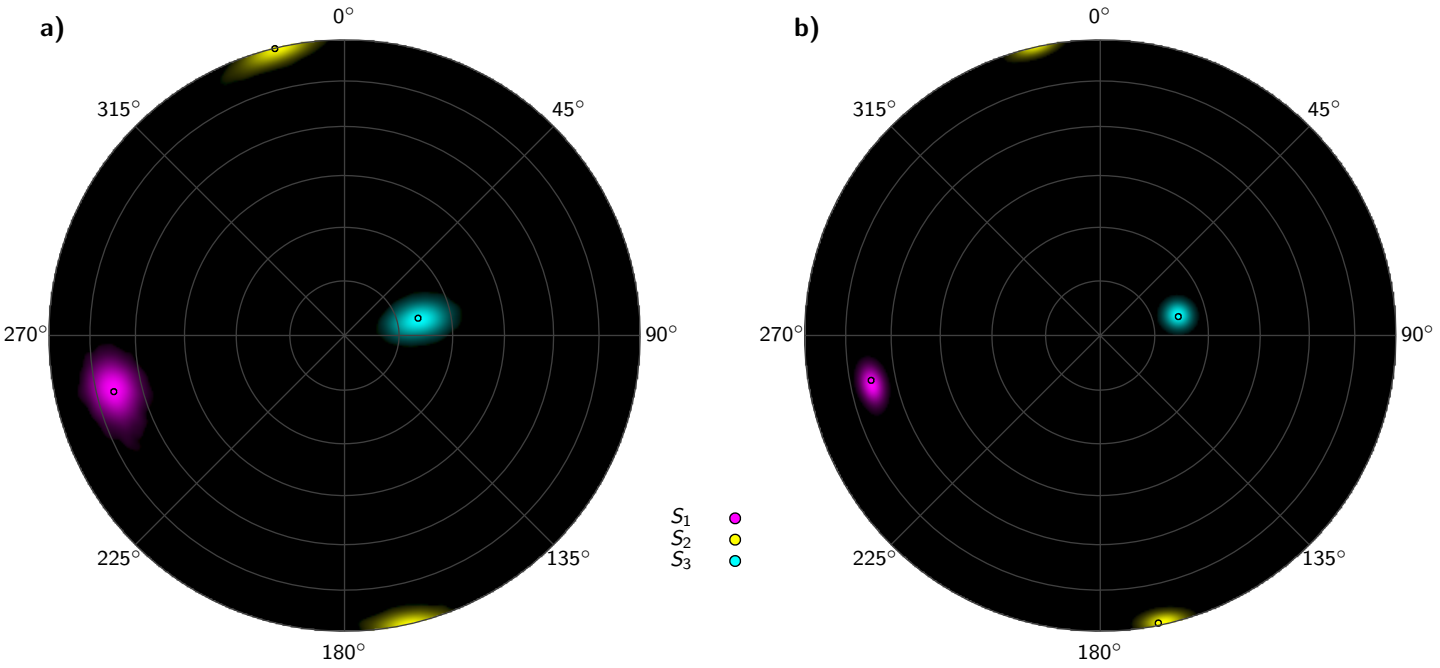


Figure 4.4: Results of conventional stress tensor inversion (a) and stress tensor inversion using weights based on ACE (b) for northern Chile. For the conventional inversion we used all events from the northern Chile (region extent in Figure 4.3d) with a maximum depth of 50 km. For the weighted inversion we used probabilities of the events belong to interface subpopulation (cyan colored in Fig. 4.3). The plots show the principal stress axes orientations (dots) in a lower hemisphere using the Lambert azimuthal equal-area projection (Schmidt net). Uncertainty distributions are estimated by bootstrapping (20,000 times) and are shown as a kernel density estimator based on the von Mises-Fisher distribution (κ=2500). Stress orientations are nearly identical for both inversions, though uncertainties are significantly reduced in the weighted inversion (b). This reduction is achieved by excluding events not associated with interface activity. Those excluded events are more likely associated with a different local stress field (e.g. Andes thrustbelt, reactivation of faults in the oceanic plate).

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

4.2 Nazca Plate

The second example covers the focal mechanisms of the GCMT catalog from 1976 - 2016 for the entire Nazca plate with eastward motion, thus this example does not cover data from the northern plate boundary due to the plate's north-south motion. As the previous example is a subset of this data set, we expect similar results for normal and reverse faulting. The bulk of the Nazca plate can be considered as a homogeneous block moving eastwards surrounded by the Cocos plate to the north, the Pacific plate to the west, the Antarctic Plate to the south and the South American Plate to the east. The catalog for this region contains 3528 events (7056 nodal planes) and the setup for ACE is as in the previous example.

The optimal mixture model (Fig. 4.7a) shows four clusters similar to those found for Northern Chile: two for reverse and normal faulting, respectively. In addition, eight clusters with rakes close to either 0° or ±180° are identified as well, indicating strike-slip. In total, five subpopulations are identified: reverse, normal, two for strike-slip (right- and left-lateral), and one for unclassified events (Fig. 4.7a). The normal and reverse subpopulations appear similar in shape and location as in case of the Northern Chile data (which are a subset here), suggesting that the stress orientations remain fairly constant along the entire west coast of South America.

The rotation angle distributions for each event type subpopulation are shown in Fig. 4.5. The ECDF for reverse faulting events (cyan) is comparable to the one from Northern Chile in Fig. 4.2, though the concentration κ is lower. This is expected due to the much larger catalog of events spanning a larger area. The normal faulting subpopulation is less well described by a rotational Fisher-Bingham distribution and the ECDF deviates from the model. This can be explained by the large variety of normal faulting events along the Nazca plate boundaries. While events to the East are mostly associated with intraslab activity, normal faulting events along the southern and western boundaries occur along the mid-ocean ridges.

A similar behavior of the rotation angle distribution is observed for strike-slip events (Fig. 4.5, green curve). This strike-slip subpopulation is associated with events along the East Pacific Rise (Fig. 4.7d, west boundary of the Nazca Plate) and events in this region are usually left-lateral. Yet another group in this subpopulation occurs along the Liquiñe-Ofqui fault in Chile (Fig. 4.7d, green colored events, south-eastern corner). This fault on the South American Plate is subject to the continental stresses there and only indirectly connected to the Nazca Plate. These shallow events along this fault are right-lateral (*Hauser*, 1991). Therefore, this subpopulation is relatively diverse, which explains the relatively low concentration κ and the deviation from the rotational Fisher-Bingham distribution.

On the other hand, strike-slip events along the southern

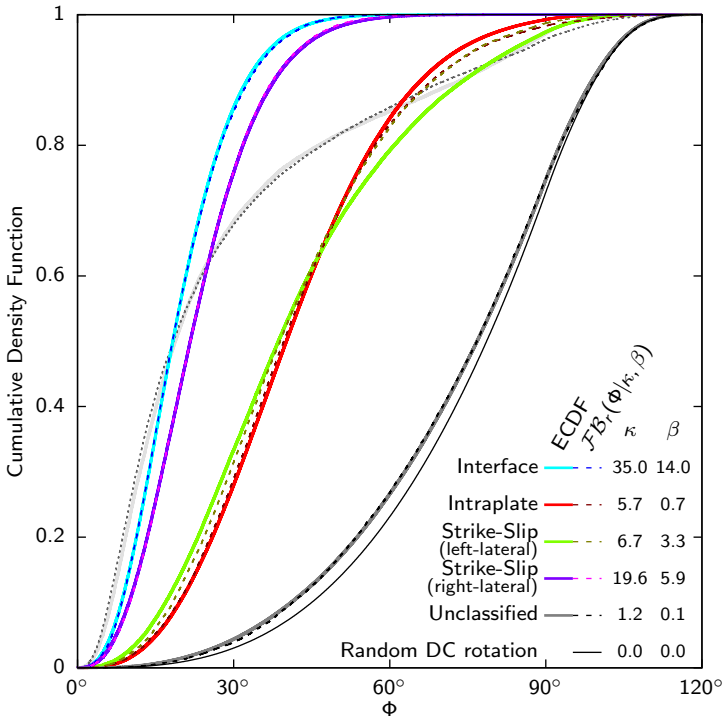


Figure 4.5: Cumulative distribution functions of the rotation angle between pairs of FMS for the Nazca Plate. The colored solid curves are estimated from the FMS data by using the subpopulation probabilities as weights. The empirical distributions are well described by rotational Fisher-Bingham distributions ($\mathcal{FB}_r$, dashed lines, parameters are in table). As for the northern Chile FMS (which are completely included in this data set), the angle distribution for unclassified FMS is nearly random and is very similar to the pure random DC distribution (thin black line). The solid light gray line shows the rotation angle distribution for pairs of FMS shallower than 100 km and with a hypocenters separated by less than 50 km. This unweighted distribution is best described by a rotational Cauchy distribution (κ_c=0.07, dotted line).

plate boundary to the Antarctic Plate are very similar to each other as shown by relatively small rotation angles (high concentration). These events are all left-lateral strike-slip. This applies also to the events further north close to the East Pacific Rise.

The derived orientation of both strike-slip subpopulations is in agreement with the plate boundaries and the plate movement directions (*Bird*, 2003) (Fig. 4.7e).

The distribution of unweighted FMS (Fig. 4.5, light gray line) is determined as for the Northern Chile example: only events with hypocenters shallower than 100 km, and distances between hypocenters of less than 50 km. The resulting distribution is similar to the one for Northern Chile (Fig. 4.5, light gray line), yet less steep due to the much larger spatial extent of the region and therefore less similar events are more likely the catalog. As for Northern Chile, the rotational Cauchy distribution fits well to empirical distribution (Fig.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

4.5, dotted gray line).

Due to the large extent of the region, we conducted stress tensor inversion with SATSI. As for northern Chile we focus on events associated with interface activity, i.e. with a depth of maximum 50 km. We subdivide the region between -85° and -60° longitude into seven strips of 5° width between latitudes -40° and -5°. Zones are shown as red horizontal lines in Fig. 4.6. We applied SATSI both conventionally without weighting and with ACE based weights. The dependency strength in Eq. 3.12 is set for both cases to $\epsilon = 0.2$ in order to minimize the trade-off between model length and variance. Results for both inversions are shown in Fig. 4.6. For each bin, we show the orientations of the principal stress axes in a lower hemispheric plot (large circles). The left column shows principal stresses for the conventional inversion and the right column for the ACE based inversion. All results are similar with the maximum principal stress S_1 pointing westward, the intermediate principal stress S_2 pointing northward (southward), and the minimum principal pointing downward and slightly eastward (depicted by the shortest arrow).

Results of the conventional inversion tend to be more tilted than results from the ACE based inversion. This is shown by the flipping of the northward/southward pointing stress axes in the ACE based inversion (Fig. 4.6, right column). These stress axes are nearly horizontal, hence they tend to flip from north to south for different tensor solution. This is not the case for the conventionally derived tensors, where all intermediate stress axes point northward.

Uncertainties are determined on the same basis as for northern Chile by bootstrapping. Uncertainties for each tensor between the conventional and ACE based inversion procedure are similar (Fig. 4.8). Individual solutions are shown by different shades of hue, with S_1 in red/purple, S_2 in orange/green, and S_3 in green/blue. However, the scatter of the tensor solutions themselves are larger for the conventional inversion. This is best demonstrated by the orientation of S_3: While for the conventional analysis each tensor uncertainty distribution is distinctively visible (Fig. 4.8a, green/blue areas), it is not the case for the ACE based inversion. Here all uncertainty distributions are in the same location (Fig. 4.8b, blue area). A similar behavior can also be observed for the maximum principal stress orientation (S_1, red/purple), and intermediate stress orientations (S_2, orange/green). For the intermediate stress orientations it also visible in Fig. 4.8 that orientations are closer to horizontal (i.e. at the edge of the plot) for the ACE based solutions than for the conventional inversion.

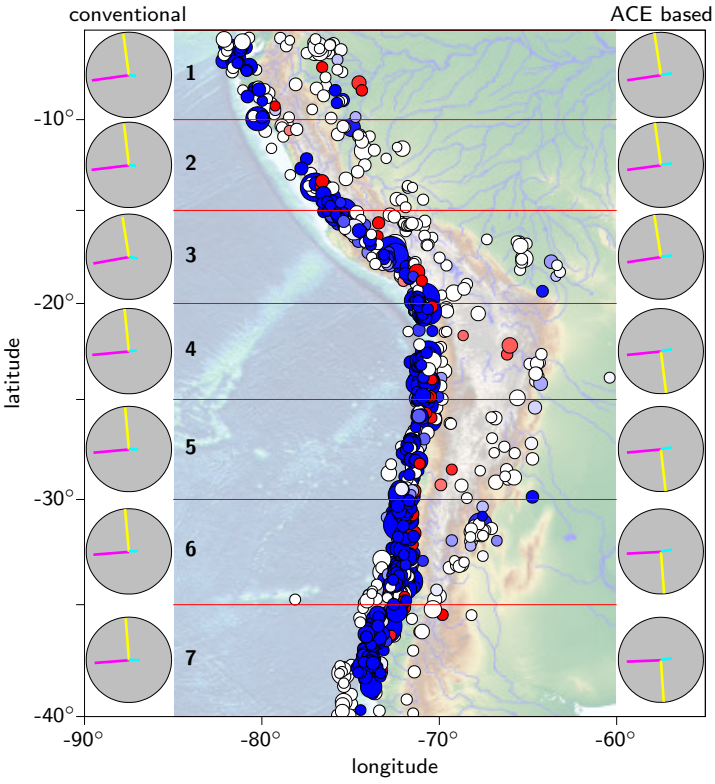
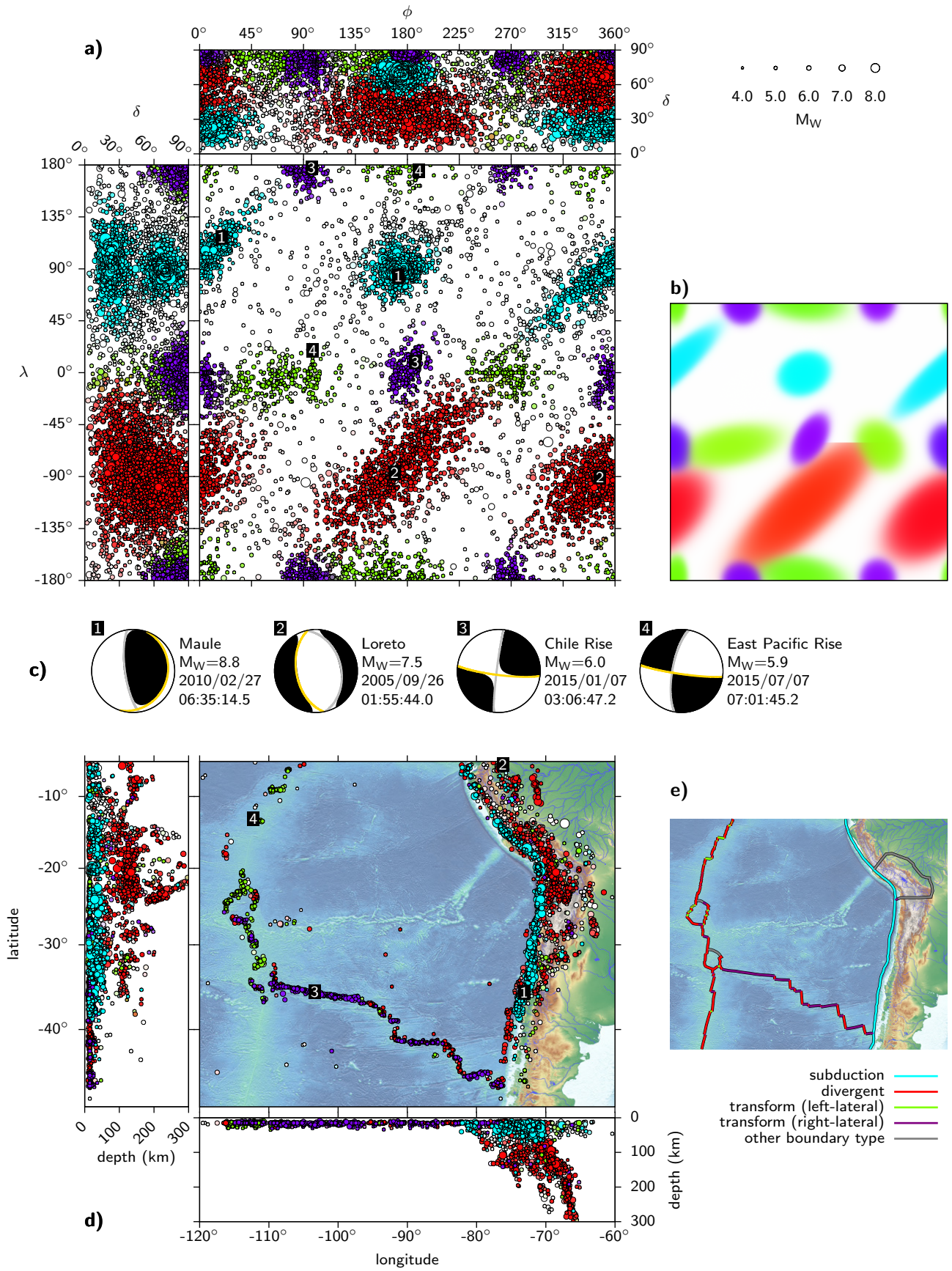


Figure 4.6: Results of conventional stress tensor inversion (left column) and stress tensor inversion using weights based on ACE (right column) by using SATSI for the South American west coast. Each tensor represents the stress in the map area next to it (areas separated along red lines). For the conventional inversion we used all events along the coast (region extent in map) with a maximum depth of 50 km. This encompasses all events in blue and white. For the ACE based inversions we used all events and the color saturation corresponds to the weight size: the more saturated, the larger the weight. The separation into blue and red colors is only necessary to indicate those events with depths less than 50 km (blue) and therefore also included in the conventional inversion (without weighting) and depths more than 50 km (red) and hence excluded from the conventional inversion. In case of the deeper events we show only events that highly probable of interface type. The stress orientation plots show principal stress orientations as lines (S_1 magenta, S_2 yellow, S_3 cyan). Projection is lower hemispheric Lambert azimuthal equal-area. Using the lower hemisphere only results in the flips of the S_2-axes (yellow) in the right column.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012



STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Figure 4.7: (left page) a) Optimal mixture model for FMS data from the Nazca plate from 1976-2016. The twelve identified clusters fall into four subpopulations: reverse (cyan), normal (red), strike-slip (chartreuse, violet). b) The reverse and normal faulting clusters are similar to the clusters shown in Fig. 4.3a,b and represent interface (cyan) and intraslab (red) events. In addition, the red clusters also include events along the mid-ocean ridges in the South and West of the Nazca Plate. Eight clusters for strike-slip are identified falling into two subpopulations. c) Four example beach ball plots for each SoF. The rupture plane and auxiliary planes are indicated as yellow and gray great-circles, respectively. The numbers correspond to the numbers in the strike-rake plane (a) and the map (d). d) Map of the Nazca plate showing the hypocenters of the GCMT catalog. Nearly all events are at the plate boundaries and the colored hypocenters of the near surface events delineate the plate boundaries as shown in (e) (*Bird*, 2003). Map data: ETOPO1 (*Amante and Eakins*, 2009). From the plate movement and plate boundaries it is possible to infer the strike-slip type: Nearly all strike-slip events along the Chile Rise (southern plate boundary with the Antarctic Plate) and along the southern East Pacific Rise form the subpopulation shown in violet. Due to the eastward motion of the Nazca Plate it is reasonable to assume these events as right-lateral. The strike-slip events shown in chartreuse are mostly along the East Pacific Rise and with left-lateral transform fault segments. A small group of this subpopulation is also located on the South American continent along the Chilean coast (lower right corner). These events are most likely associated with the right-lateral Liquiñe Ofqui fault.

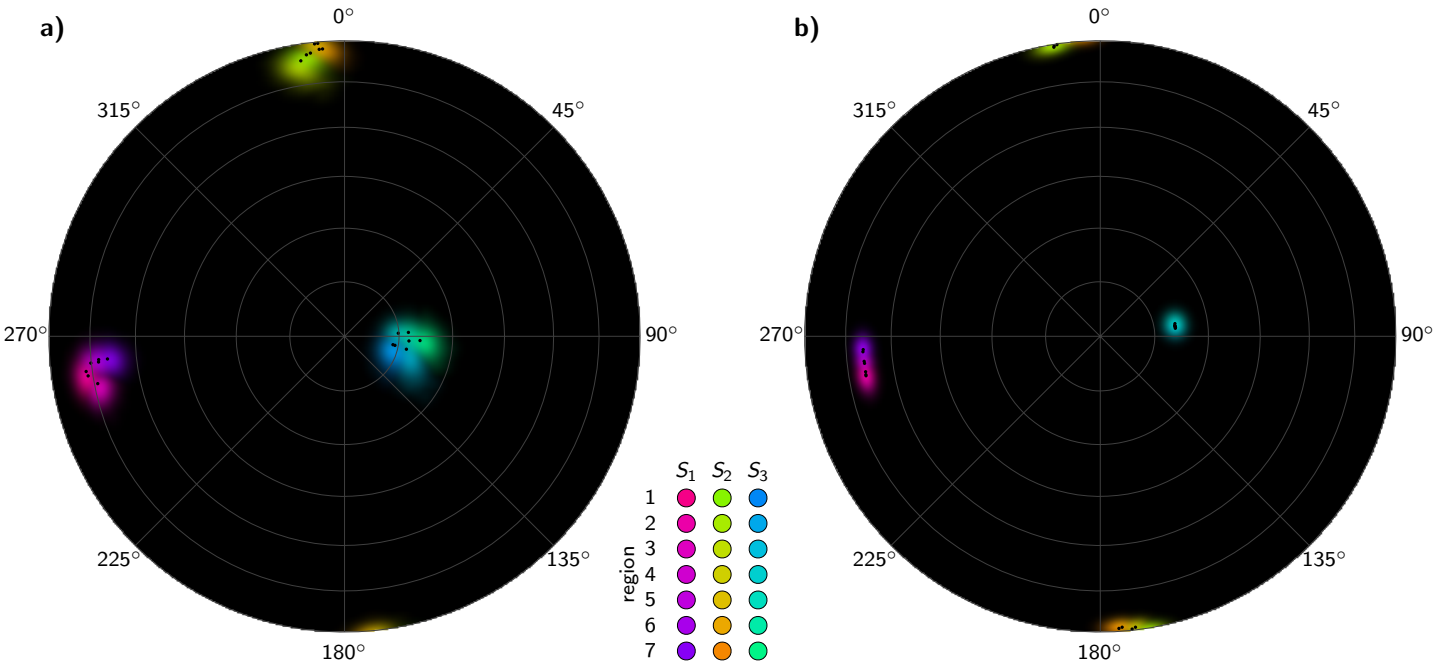


Figure 4.8: Same as Fig. 4.4 but for results in Fig. 4.6 (South American west coast) of SATSI based stress tensor inversion. The different shades for each stress axis distribution represent the different tensors for each region (base colors and numbers correspond to axes colors and numbers in Fig. 4.6). Uncertainties (shown as colored KDE) for each tensor are smaller for the ACE based inversion compared to the conventional weighting. In addition, results on the left are more scattered, e.g. S_1 orientations (black dots) show an unsorted "C" shaped pattern. On the hand, ACE based results show steady increase of the S_1 azimuth, while the dips remain constant. The orientation of the seven S_3-axes of the ACE based inversion (right) are within the uncertainties of each individual tensor, thus resulting in mixing the color shades in the plot. Also the S_2-axes are closer to horizontal for the ACE based inversion, resulting in distributions in the upper and lower half of the plot which explains the flip of the yellow axes in Fig. 4.6.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

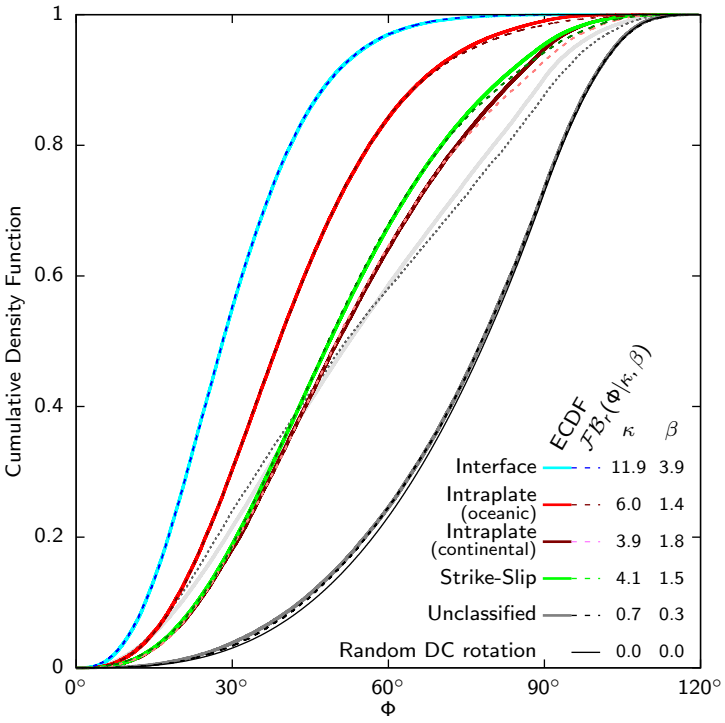


Figure 4.9: Cumulative distribution functions of the rotation angle between pairs of FMS from Kyūshū. The colored solid curves are estimated from the FMS data by using the subpopulation probabilities as weights. The empirical distributions are well described by rotational Fisher-Bingham distributions ($\mathcal{FB}_r$, dashed lines, parameters are in table). The angle distribution for unclassified FMS is nearly random since it is close to the pure random DC distribution (thin black line). The solid light gray line shows the rotation angle distribution for pairs of FMS shallower than 100 km and with a hypocenters separated by less than 50 km. This unweighted distribution is best described by a rotational Cauchy distribution (κ_c=0.23, dotted line). The distribution of the rotation angle indicates more tectonic heterogeneity compared to the previous examples from South America (Fig. 4.2, 4.5).

Figure 4.10: (right page) a) Optimal mixture model for FMS data from Kyūshū (Japan) from 1997-2016. The ten identified clusters fall into four subpopulations: reverse (cyan), two normal (red) and strike-slip (green). The two normal fault subpopulations show some differences: The clusters are separated by appr. 90° of strike, i.e. irrespective of the nodal plane, the rupture planes are perpendicular to each other. For the subpopulation shown in dark red the dips of both nodal plane clusters are appr. the same, while for the subpopulation in bright red the dips differ significantly for both clusters. The difference of dips in the clusters is also observable for the data from South America (Fig. 4.3a, 4.7a). c) Map of hypocenter locations shows most reverse faulting events along the eastern part of Kyūshū along the interface of the Philippine Plate which is subducted from East under the Eurasian Plate. Normal faulting FMS shown in bright red are close to interface related events and are associated with activity within the oceanic Philippine Plate. The normal faulting events in dark red are located further away from the subduction zone, relatively shallow and close to the strike-slip events. In central Kyūshū where the 2016 Kumamoto earthquake sequence occurred (moment tensor of main shock shown in (d)), events of normal and strike-slip subpopulations are spatially close. Events in this region may form a continuous transition from one SoF to another which is also indicated by the cluster alignment in (a).

4.3 Kyūshū (SW Japan)

For the third example we investigated focal mechanisms of the NIED catalog from 1997 - 2016 for the island of Kyūshū and surroundings in southwestern Japan. Kyūshū is located on the Eurasian plate, west of the Philippine Plate. The island is traversed by the Futagawa and Hinagu fault systems from Southwest to Northeast. The April 2016 Kumamoto earthquake sequence (main shock: M_W 7.1) occurred on this fault system (*Kubo et al.*, 2016). The catalog contains 1741 events (3482 nodal planes).

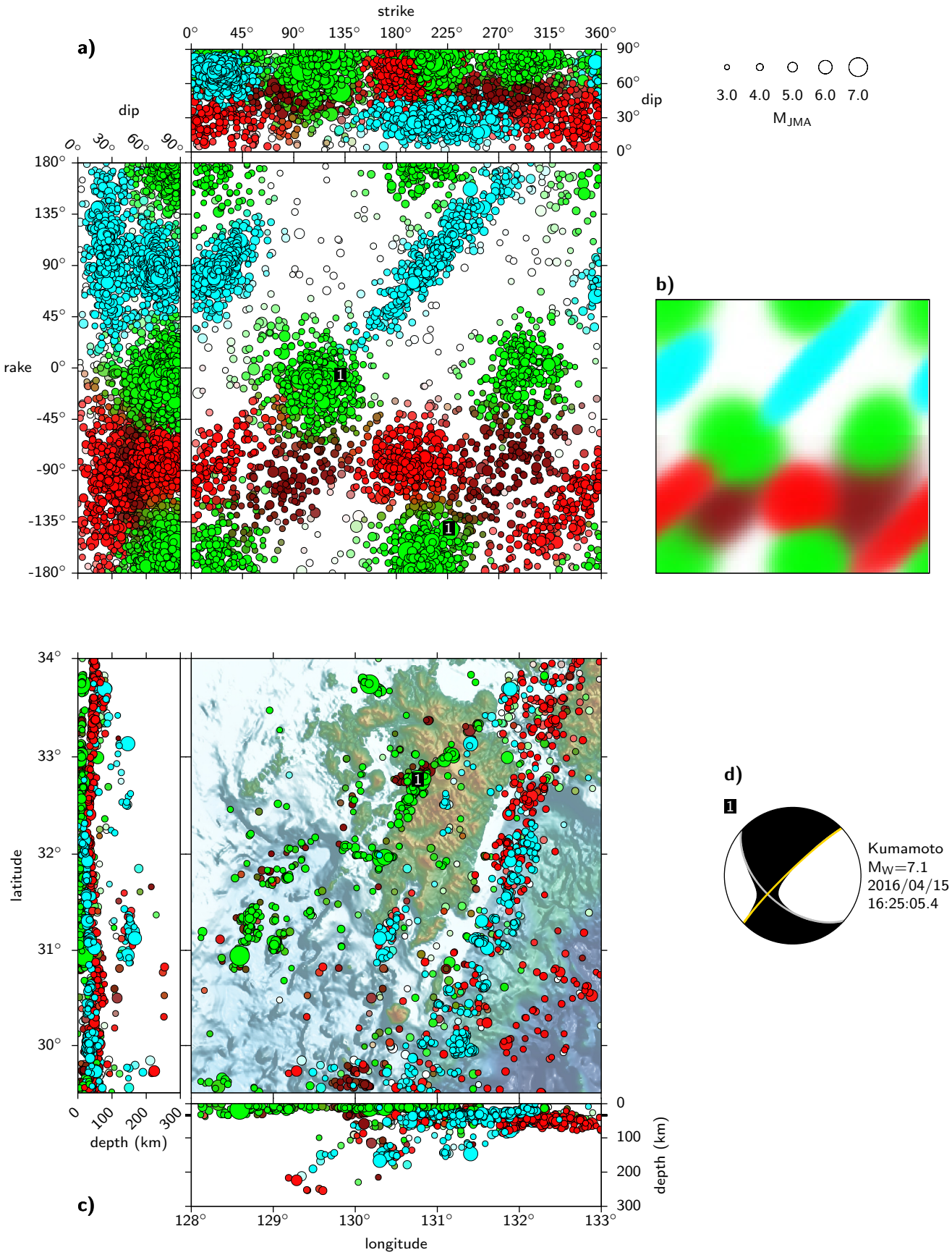
The optimal mixture model (Fig. 4.10a) identifies eleven clusters: one for unclassified faulting, two for reverse faulting, and four for normal and strike-slip faulting, respectively.

The two different types of normal faulting are also spatially separated (Fig. 4.10d). The subpopulation of normal faulting events with different mean dips of its two clusters (Fig. 4.10a, bright red) are mostly located in the slab of the subducted Philippine plate, while the other subpopulation of normal events is located in the crust of Kyūshū (Fig. 4.10a, dark red).

The rotation angle distribution are less concentrated compared to the results of the Nazca plate, even for reverse faulting (Fig. 4.9, cyan line). Though all subpopulations are well described by the rotational Fisher-Bingham distribution, both continental intraplate ((Fig. 4.9, dark red) and strike-slip (green line) events show deviations from the rotational Fisher-Bingham distribution and exhibit a slight jump at appr. 90°. This jump is similar to the rotational Cauchy distribution for large rotation angles, though the bump here is smaller in size. This fact and the relatively low concentration κ might be an indicator for a comparatively heterogeneous setting. The unclassified events are almost purely randomly distributed FMS.

The empirical distribution with unweighted data is determined as in the previous examples by considering only events shallower than 100 km and a maximum distance of 50 km between hypocenters. Similarly to the relatively low concentrations of the rotational Fisher-Bingham distributions, the unweighted rotation angle distribution demonstrates a heterogeneous setting for Kyūshū, as nearly half of the FMS pairs have rotation angles of more than 60° (Fig. 4.9, light gray line). Irrespective of the complexity, the rotation angle distribution is well described by the rotational Cauchy distribution (Fig. 4.9, dotted gray line).

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012



STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Chapter 5

Discussion & Conclusion

We introduced ACE as a purely data driven algorithm to identify clusters of earthquakes, solely based on a focal mechanism's strike, rake, and dip. The results ACE provides are in agreement with the general knowledge about the regions investigated. Within a region with relatively constant stress orientations, hypocenters are not directly required to classify the data. Therefore, classification errors are reduced, compared to classical deterministic classification schemes as presented by *García et al.* (2012). We also successfully demonstrated that the clusters improve estimates of stress tensors and can be further used for nodal plane disambiguation in the presence of geological/tectonic data.

Few strike-slip are also shown at larger depths (> 50 km, Fig. 4.7d, left and bottom). This misclassification can be caused by the fact, that certain more complex rupture processes are insufficiently represented by a FMS. Another explanation can be found by the fact of an insufficient amount of data. If a particular rupture process is represented by a small number of events, then it is unlikely that this particular process is assigned to its own cluster.

The automatic classification can be used for a wide range applications in the seismology and stress communities, e.g. for seismic hazard assessment, selection of existing ground motion prediction equations (GMPE's), selection of earthquakes to derive new GMPE's, SoF-specifc b-value estimation, selection of focal mechanism as stress indicators (World Stress Map Project) etc.

ACE - in conjunction with geological databases - can also be applied in rapid earthquake assessment to help identify the rupture plane of a newly calculated FMS. In this case, an existing FMS catalog is used as a training data set to derive a mixture model, which is representative for a region of interest. When a new FMS datum is available, the probability of the SoF can be calculated, and if additional geological data is available, the rupture plane orientation can be derived.

We showed the application of the mixture model probabilities of ACE as weights for the determination of the rotation angle distribution of focal mechanisms. The resulting weighted empirical distributions are well described by rotational Fisher-Bingham distributions even in tectonically complex settings. This provides an additional perspective on the distribution of FMS because the weighted distributions show high consistency even to spatially distant events. This effectively reduces the constraint of spatial closeness when considering unweighted data (*Kagan*, 1992; *Kagan and Jackson*, 2015). The investigation of the distributions of rotation angle between FMS pairs with ACE based weights also showed the separability of the data into rotationally consistent subsets and a nearly purely randomly orientated FMS subset (the noise/unclassified subpopulation). The former can be associated with large scale tectonic features (e.g. plate interface, subducting slab). The latter subset is only "nearly" purely random as it is not possible to separate those events from a consistent subpopulation (e.g. interface related) that are only by chance similar to the subpopulation's events. This is the main reason why the rotation angle distributions of the unclassified subpopulations are not completely following the random DC rotation.

Stress tensor inversion results with ACE based weighting of the FMS improve stress tensor uncertainties considerably. However, this reduction in uncertainties affects the precision of the results, not necessarily their accuracy. If only one plane of a conjugate fault system is activated (as usually is the case of subduction zones since subduction does not occur on conjugate planes), then the resulting stress tensor can be biased. The bias affects the minimum and maximum principal stress directions and has a maximum magnitude of half the (usually unknown) angle of internal friction in the plane of the maximum and minimum stress orientations. If only one conjugate plane is activated, then the inversion will converge to solutions identical to the PBT axes. This convergence occurs irrespective of a priori knowledge of the rupture and auxiliary plane orientations.

The results of ACE based stress tensor inversion with nodal plane disambiguation are useful for rupture plane distance calculation (with magnitude-fault size relations from e.g. *Strasser et al.* (2010)), an important distance metric for ground motion prediction equations.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Appendix A

Mathematical concepts

A.1 Derivation of angular derivatives

A.1.1 Relations for rupture plane

The equations for the slip and normal vectors in terms of strike (ϕ), rake (λ) and dip (δ) are given by

$$\hat{\mathbf{n}} = \begin{pmatrix} -\sin\delta\sin\phi \\ -\sin\delta\cos\phi \\ \cos\delta \end{pmatrix} \tag{A.1}$$

$$\hat{\mathbf{d}} = \begin{pmatrix} \sin\lambda\cos\delta\sin\phi + \cos\lambda\cos\phi \\ \sin\lambda\cos\delta\cos\phi - \cos\lambda\sin\phi \\ \sin\lambda\sin\delta \end{pmatrix} \tag{A.2}$$

Let $\hat{\mathbf{d}}_2$ denote the slip vector of the auxiliary and $\hat{\mathbf{n}}_1$ the normal vector of the rupture plane. Since the nodal planes are orthogonal to each other, the vectors are identical (e.g. *Stein and Wysession*, 2003)

$$\hat{\mathbf{d}}_2 = \hat{\mathbf{n}}_1 \quad \text{or} \quad \hat{\mathbf{d}}_1 = \hat{\mathbf{n}}_2 \tag{A.3}$$

Following the Wallace-Bott hypopthesis, the slip and normal vectors are in the same plane as the maximum and minimum principal stress axes. This plane's normal vector is parallel to both the intermediate principal stress axis and the cross product of the slip and normal vector:

$$\hat{\mathbf{b}} = \hat{\mathbf{n}} \times \hat{\mathbf{d}} \tag{A.4}$$

$$= \begin{pmatrix} \cos\lambda\cos\delta\sin\phi - \sin\lambda\cos\phi \\ \cos\lambda\cos\delta\cos\phi + \sin\lambda\sin\phi \\ \cos\lambda\sin\delta \end{pmatrix}. \tag{A.5}$$

With the cross product (and intermediate stress axis) fixed, it follows

$$\hat{\mathbf{b}} = \text{const.} \tag{A.6}$$

Now, let the intermediate stress vector be defined as $\hat{\mathbf{b}}(\lambda(\phi,\delta),\delta,\phi)$ and

$$\frac{\partial\hat{\mathbf{b}}}{\partial\phi} = 0 \tag{A.7}$$

And with Eq. A.2:

$$\begin{aligned} 0 = &-\sin\lambda\frac{\partial\lambda}{\partial\phi}\cos\delta\sin\phi - \cos\lambda\cos\delta\cos\phi \\ &-\cos\lambda\frac{\partial\lambda}{\partial\phi}\cos\phi + \sin\lambda\sin\phi \end{aligned} \tag{A.8}$$

$$\begin{aligned} 0 = &-\sin\lambda\frac{\partial\lambda}{\partial\phi}\cos\delta\cos\phi - \cos\lambda\cos\delta\sin\phi \\ &+\cos\lambda\frac{\partial\lambda}{\partial\phi}\sin\phi + \sin\lambda\cos\phi \end{aligned} \tag{A.9}$$

Multiplying the first line with $\cos\phi$ and the second with $\sin\phi$ and subtracting the second from the first equation reduces to the relation

$$\frac{\partial\lambda}{\partial\phi} = \cos\delta \tag{A.10}$$

The derivative with respect to δ:

$$\frac{\partial\hat{\mathbf{b}}}{\partial\delta} = 0 \tag{A.11}$$

And with Eq. A.2:

$$\begin{aligned} 0 = &-\sin\lambda\frac{\partial\lambda}{\partial\delta}\cos\delta\sin\phi - \cos\lambda\sin\delta\sin\phi \\ &-\cos\lambda\frac{\partial\lambda}{\partial\delta}\cos\phi \end{aligned} \tag{A.12}$$

$$\begin{aligned} 0 = &-\sin\lambda\frac{\partial\lambda}{\partial\phi}\cos\delta\cos\phi - \cos\lambda\sin\delta\cos\phi \\ &+\cos\lambda\frac{\partial\lambda}{\partial\delta}\sin\phi \end{aligned} \tag{A.13}$$

Again, multiplying the first line with $\cos\phi$ and the second with $\sin\phi$ and subtracting the second from the first equation reduces to the relation

$$\cos\lambda\frac{\partial\lambda}{\partial\delta} = 0 \tag{A.14}$$

The derivatives in Eq. A.10 and A.14 are also found by setting $\hat{\mathbf{d}} = \text{const.}$

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

A.2 Wrapped Normal Distribution

The bivariate normal distribution of a vector $\boldsymbol{\upsilon} = (\upsilon_1, \upsilon_2)$:

$$\mathcal{N}(\boldsymbol{\upsilon}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi\sqrt{\det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{\upsilon} - \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\upsilon} - \boldsymbol{\mu})^T\right) \tag{A.15}$$

where $\boldsymbol{\mu} = (\mu, \nu)$ is the mean vector and $\boldsymbol{\Sigma}$ the covariance matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}, \tag{A.16}$$

with σ^2 and τ^2 as the variances and ρ is the (population) correlation coefficient.

The bivariate wrapped normal (WN) distribution is

$$\mathcal{N}_w(\phi, \lambda|\mu, \nu, \sigma^2, \tau^2, r) = \sum_{u,v\in\mathbb{Z}} \mathcal{N}(\phi + 2\pi u, \lambda + 2\pi v|\mu, \nu, \sigma^2, \tau^2, r) \tag{A.17}$$

The strike ϕ is in the range $[0, 2\pi]$ and has mean μ and variance σ^2. The rake λ is in the range $[-\pi, \pi]$ with mean ν and variance τ^2. The estimators of these parameters are found by maximum likelihood. Since the parameters of strike and rake have similar estimators, let $\boldsymbol{\theta} = (\phi, \lambda)$ represent the data, $\boldsymbol{\mu} = (\mu, \nu)$ the means and $\boldsymbol{\sigma} = (\sigma^2, \tau^2)$ the variances.

In order to address the circular property of $\boldsymbol{\theta}$, which is also inherent in the parameters of the WN distribution, a change of variable is performed:

$$\boldsymbol{\Theta} = e^{i\boldsymbol{\theta}}. \tag{A.18}$$

The mean is found by

$$\bar{\boldsymbol{\Theta}} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{\Theta}_n, \tag{A.19}$$

and the maximum likelihood estimate of $\boldsymbol{\mu}$ of the WN distribution is

$$\hat{\boldsymbol{\mu}} = \arg \bar{\boldsymbol{\Theta}}. \tag{A.20}$$

The quadratic length of the mean vector, given by $\bar{\boldsymbol{\Theta}}\bar{\boldsymbol{\Theta}}^*$ ($*$ denotes complex conjugate), is used to estimate the variances of $\hat{\boldsymbol{\sigma}}$

$$\hat{\boldsymbol{\sigma}} = -\ln\left(\frac{N}{N-1}\left(\bar{\boldsymbol{\Theta}}\bar{\boldsymbol{\Theta}}^* - \frac{1}{N}\right)\right). \tag{A.21}$$

For large N it is sufficient to approximate the variance by

$$\hat{\boldsymbol{\sigma}} = -\ln(\bar{\boldsymbol{\Theta}}\bar{\boldsymbol{\Theta}}^*). \tag{A.22}$$

When applied within the expectation-maximization algorithm, this (biased) estimator must be used as the sample size is lost when the data are weighted according to the distribution functions.

With the mean estimated by Eq. A.19, the circular sample correlation coefficient is calculated by

$$r = \frac{\sum_{n=1}^{N} \sin(\phi_n - \hat{\mu}) \sin(\lambda_n - \hat{\nu})}{\sqrt{\sum_{n=1}^{N} \sin^2(\phi_n - \hat{\mu})}\sqrt{\sum_{n=1}^{N} \sin^2(\lambda_n - \hat{\nu})}}. \tag{A.23}$$

A.3 Beta Distribution

Similar to the wrapped normal distribution, the beta distribution is chosen for its wide range of shapes (other distributions are special/limiting cases of the beta distribution), its finite support, and the availability of simple parameter estimators.

The generalized beta distribution in an arbitrary interval (a, b) is given by

$$\mathcal{B}(x|\alpha, \beta, a, b) = \frac{(x-a)^{\alpha-1}(b-x)^{\beta-1}}{(b-a)^{\alpha+\beta-1}B(\alpha, \beta)}, \tag{A.24}$$

where α and β are the shape parameters. The beta function $B(\cdot, \cdot)$ is the normalization constant and is given by

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \tag{A.25}$$

with $\Gamma(\cdot)$ as the gamma function. When programming the above equation, it as advisable to define it terms of its logarithms by using the log-gamma function (e.g. `lgamma` in C/C++, `gammaln` in MATLAB®are provided])/Python™. As the gamma function increases to very large values even for small arguments, summing and subtracting the logarithms and exponentiation of the result at the end prevents arithmetic overflow.

The beta distribution has finite support in the range of (a, b), i.e. following the definition of *Johnson et al.* (1995), and leaves out the interval limits, due to the singularities at $\{0, 1\}$ for $0 < \alpha < 1$ or $0 < \beta < 1$. When applied to data, i.e. the dip angle δ, the interval limits must be defined to include all values. In case of the dip δ, the interval is set to $(-1°, 91°)$.

The estimators of the shape parameters are based on the method of moments. Only the mean m_δ and variance v_δ of the dip and their interval normalized counterparts m and v, respectively, are required:

$$m = \frac{m_\delta - a}{b - a} \qquad m_\delta = \frac{1}{N}\sum_{i=1}^{N} \delta_i \tag{A.26}$$

$$v = \frac{v_\delta}{(b-a)^2} \qquad v_\delta = \frac{1}{N-1}\sum_{i=1}^{N} (\delta_i - m_\delta)^2 \tag{A.27}$$

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

The shape parameters are found from the definitions of the moments in terms of the shape parameters

$$m = \frac{\alpha}{\alpha+\beta} \qquad v = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}, \tag{A.28}$$

and solving for the shape parameters results in

$$\hat{\alpha} = m\left(\frac{m(1-m)}{v} - 1\right) \tag{A.29}$$

$$\hat{\beta} = (1-m)\left(\frac{m(1-m)}{v} - 1\right). \tag{A.30}$$

These relations hold for $v < m(1-m)$. In few instances, this inequality may not hold within the EM-algorithm. In order to avoid discarding clusters with ill-conditioned shape parameters, we set $\hat{\alpha} = \hat{\beta} = 1$, i.e. the dip is then described by a uniform distribution. Note, that this overriding rule may become obsolete again as the iterations proceed and the distributions change.

A.4 Change of variable

The change of variables of a multivariate probability density function from $\mathbf{x}$ to $\mathbf{y}$ is given by

$$P(\mathbf{y}) = P(\mathbf{x})|\det \mathbf{J}|, \tag{A.31}$$

where $\mathbf{J}$ is the Jacobian matrix. The determinant of the Jacobian for the nodal plane angles is

$$\det(\mathbf{J_1}) = \begin{vmatrix} \frac{\partial\phi_1}{\partial\phi_2} & \frac{\partial\phi_1}{\partial\lambda_2} & \frac{\partial\phi_1}{\partial\delta_2} \\ \frac{\partial\lambda_1}{\partial\phi_2} & \frac{\partial\lambda_1}{\partial\lambda_2} & \frac{\partial\lambda_1}{\partial\delta_2} \\ \frac{\partial\delta_1}{\partial\phi_2} & \frac{\partial\delta_1}{\partial\lambda_2} & \frac{\partial\delta_1}{\partial\delta_2}. \end{vmatrix} \tag{A.32}$$

Solving $\hat{\mathbf{d}}_1 = \hat{\mathbf{n}}_2$ for ϕ_1, λ_1, and δ_1 in terms of ϕ_2, λ_2, and δ_2 (Eq. A.1 and A.2) gives following three equations:

$$\phi_1 = \phi_2 - \arcsin\left(\frac{\cos\lambda_2}{\sqrt{1-\sin^2\lambda_2\sin^2\delta_2}}\right) \tag{A.33}$$

$$\lambda_1 = \arcsin\left(\frac{\cos\delta_2}{\sqrt{1-\sin^2\lambda_2\sin^2\delta_2}}\right) \tag{A.34}$$

$$\delta_1 = \arcsin\sqrt{1-\sin^2\lambda_2\sin^2\delta_1} \tag{A.35}$$

From equations A.33, A.34 and A.35, it can be readily seen that

$$\frac{\partial\phi_1}{\partial\phi_2} = 1,\ \frac{\partial\lambda_1}{\partial\phi_2} = 0,\ \text{and}\ \frac{\partial\delta_1}{\partial\phi_2} = 0, \tag{A.36}$$

and the determinant of the Jacobian matrix reduces to

$$\det(\mathbf{J_1}) = \frac{\partial\lambda_1}{\partial\lambda_2}\frac{\partial\delta_1}{\partial\delta_2} - \frac{\partial\lambda_1}{\partial\delta_2}\frac{\partial\delta_1}{\partial\lambda_2}. \tag{A.37}$$

A common denominator of these derivatives is

$$\sqrt{1-\sin^2\lambda_2\sin^2\delta_2} = \sin\delta_1, \tag{A.38}$$

the right hand from Eq. A.35 is used for brevity in the following equations. The derivatives are

$$\begin{aligned} \frac{\partial\lambda_1}{\partial\lambda_2} &= \frac{\cos\delta_2\sin\delta_2\sin\lambda_2}{\sin^2\delta_1} & \frac{\partial\lambda_1}{\partial\delta_2} &= -\frac{\cos\lambda_2}{\sin^2\delta_1} \\ \frac{\partial\delta_1}{\partial\lambda_2} &= \frac{\cos\lambda_2\sin\delta_2}{\sin\delta_1} & \frac{\partial\delta_1}{\partial\delta_2} &= \frac{\sin\lambda_2\cos\delta_2}{\sin\delta_1} \end{aligned} \tag{A.39}$$

and the determinant becomes

$$\det(\mathbf{J_1}) = \frac{\sin\delta_2}{\sin^3\delta_1}\left(\sin^2\lambda_2\cos^2\delta_2 + \cos^2\lambda_2\right). \tag{A.40}$$

From the Pythagorean trigonometric identity follows

$$\det(\mathbf{J_1}) = \frac{\sin\delta_2}{\sin^3\delta_1}\left(\sin^2\lambda_2\sin^2\delta_2 + 1\right), \tag{A.41}$$

and finally with Eq. A.35, the determinant of the Jacobian is

$$\det(\mathbf{J_1}) = -\frac{\sin\delta_2}{\sin\delta_1}. \tag{A.42}$$

The magnitude of the determinant is plugged into Eq. A.31 and the change of variable from a distribution given in rupture plane angles to a distribution given in auxiliary plane angles is given by

$$P(\phi_2,\lambda_2,\delta_2) = P(\phi_1,\lambda_1,\delta_1)\frac{\sin\delta_2}{\sin\delta_1}. \tag{A.43}$$

A.5 Kernel Density Estimator (KDE)

The kernel density estimator (KDE) for the nodal plane data with variance σ^2_{KDE} for the strike (ϕ) and rake (λ) strike and rake, and bandwidth b_{KDE} for the dip is the product of a bivariate normal distribution with the wrapping property for strike and rake and beta distribution for the dip. Therefore, this KDE is the non-parametric counterpart of the mixture model.

$$\begin{aligned} h(\phi,\lambda,\delta) = \frac{1}{2N\sigma_{\mathrm{KDE}}\sqrt{2\pi}}\sum_{i=1}^{2N}\frac{\delta_i^{\alpha_i-1}\left(\frac{\pi}{2}-\delta_i\right)^{\beta_i-1}}{(\frac{\pi}{2})^{\alpha_i+\beta_i-1}B(\alpha_i,\beta_i)}\times \\ \sum_{u=-1}^{1}\sum_{v=-1}^{1}\exp\Big\{-\frac{1}{2\sigma^2_{\mathrm{KDE}}} \\ \times\left[(\phi-\phi_i+2u\pi)^2+(\lambda-\lambda_i+2v\pi)^2\right]\Big\} \end{aligned} \tag{A.44}$$

The beta distribution KDE is based on *Chen* (1999). We modify the definition to apply the equation directly to the

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

support interval $(0, \frac{\pi}{2})$. Furthermore, we exclude the support limits from the beta distribution as described in appendix A.3. The parameters of the ith beta distribution, α_i and β_i, for the kernel with bandwidth b (subscript KDE dropped for brevity) in terms of dip δ_i are defined as:

$$\alpha_i = \begin{cases} \rho(\delta_i) & \text{if } \delta_i \in (0, 2b) \\ \frac{\delta_i}{b} & \text{if } \delta_i \in \left[2b, \frac{\pi}{2}\right) \end{cases} \tag{A.45}$$

$$\beta_i = \begin{cases} \frac{\frac{\pi}{2} - \delta_i}{b} & \text{if } \delta_i \in \left(0, \frac{\pi}{2} - 2b\right] \\ \rho\left(\frac{\pi}{2} - \delta_i\right) & \text{if } \delta_i \in \left(\frac{\pi}{2} - 2b, \frac{\pi}{2}\right) \end{cases} \tag{A.46}$$

where

$$\rho(x) = 2b^2 + \frac{5}{2} - \sqrt{4b^4 + 6b^2 + \frac{5}{4} - x^2 - \frac{x}{b}} \tag{A.47}$$

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Appendix B

Acknowledgments

The authors would like to thank Moritz Ziegler, Dietrich Stromeyer, and Konrad Schellbach for helpful discussions on stress, mathematics and logic. Thanks to Anne Strader for constructive comments on the manuscript. The focal mechanism data used in this study are freely available at `http://www.globalcmt.org/` (GCMT), `http://geofon.gfz-potsdam.de/` (GEOFON), `http://earthquake.usgs.gov/earthquakes/search/` (USGS), `http://www.fnet.bosai.go.jp/event/search.php?LANG=en` (NIED), `https://wwweic.eri.u-tokyo.ac.jp/tseis/junecfm2/index.html` (JUNEC FM^2) and `http://www.bo.ingv.it/RCMT/` (RCMT). Most plots are generated with gnuplot (`http://gnuplot.sourceforge.net/`), Fig. 2.3 was generated with blender (`https://www.blender.org/`). Geographic maps are generated with GMT (*Wessel et al.*, 2013), topographic data is based on ETOPO1 (*Amante and Eakins*, 2009). Sebastian Specht acknowledges support from the DFG research training group "Natural Hazards and Risks in a Changing World" (Grant No. GRK 2043/1). The code for ACE (written in C++) is accessible at `https://github.com/AgentSmith660/ACE`.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Bibliography

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Transactionson Automatic Control*, *19*(6), 716–723, http://doi.org/10.1109/TAC.1974.1100705.

Amante, C., and B. W. Eakins (2009), ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis.

Bird, P. (2003), An updated digital model of plate boundaries, *Geochemistry, Geophysics, Geosystems*, *4*(3), n/a–n/a, http://doi.org/10.1029/2001GC000252.

Bott, M. H. P. (1959), The Mechanics of Oblique Slip Faulting, *Geological Magazine*, *96*(02), 109, http://doi.org/10.1017/S0016756800059987.

Célérier, B., A. Etchecopar, F. Bergerat, P. Vergely, F. Arthaud, and P. Laurent (2012), Inferring stress from faulting: From early concepts to inverse methods, *Tectonophysics*, *581*, 206–219, http://doi.org/10.1016/j.tecto.2012.02.009.

Cesca, S., A. T. Sen, and T. Dahm (2014), Seismicity monitoring by cluster analysis of moment tensors, *Geophysical Journal International*, *196*(3), 1813–1826, http://doi.org/10.1093/gji/ggt492.

Chen, S. X. (1999), Beta kernel estimators for density functions, *Computational Statistics & Data Analysis*, *31*(2), 131–145, http://doi.org/10.1016/S0167-9473(99)00010-9.

Dempster, A. A., N. N. Laird, and D. D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society Series B Methodological*, *39*(1), 1–38, http://doi.org/10.2307/2984875.

Dupin, J.-M., W. Sassi, and J. Angelier (1993), Homogeneous stress hypothesis and actual fault slip: a distinct element analysis, *Journal of Structural Geology*, *15*(8), 1033–1043, http://doi.org/10.1016/0191-8141(93)90175-A.

Dziewonski, A. M., T.-A. Chou, and J. H. Woodhouse (1981), Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, *86*, 2825–2852.

Efron, B. (1979), Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics*, *7*(1), 1–26, http://doi.org/10.1214/aos/1176344552.

Ekström, G., M. Nettles, and A. M. Dziewoński (2012), The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Physics of the Earth and Planetary Interiors*, *200-201*, 1–9, http://doi.org/10.1016/j.pepi.2012.04.002.

Ester, M., H.-p. Kriegel, J. S, X. Xu, J. Sander, and X. Xu (1996), A density-based algorithm for discovering clusters in large spatial databases with noise, in *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96) A*, pp. 226–231, AAAI Press, http://doi.org/10.1.1.71.1980.

Feller, W. (1971), *An introduction to probability and its applications, Vol. II.*, Wiley, New York.

Fisher, N. I., and A. J. Lee (1983), A correlation coefficient for circular data, *Biometrika*, *70*(2), 327–332, http://doi.org/10.1093/biomet/70.2.327.

Frohlich, C. (1992), Triangle diagrams: ternary graphs to display similarity and diversity of earthquake focal mechanisms, *Physics of the Earth and Planetary Interiors*, *75*(1-3), 193–198, http://doi.org/10.1016/0031-9201(92)90130-N.

Frohlich, C., and K. D. Apperson (1992), Earthquake focal mechanisms, moment tensors, and the consistency of seismic activity near plate boundaries, *Tectonics*, *11*(2), 279–296, http://doi.org/10.1029/91TC02888.

García, D., D. J. Wald, and M. G. Hearne (2012), A Global Earthquake Discrimination Scheme to Optimize Ground-Motion Prediction Equation Selection, *Bulletin of the Seismological Society of America*, *102*(1), 185–203, http://doi.org/10.1785/0120110124.

Gephart, J. W., and D. W. Forsyth (1984), An improved method for determining the regional stress tensor using earthquake focal mechanism data: Application to the San Fernando Earthquake Sequence, *Journal of Geophysical Research*, *89*(B11), 9305, http://doi.org/10.1029/JB089iB11p09305.

Händel, A., S. Specht, N. M. Kühn, and F. Scherbaum (2014), Mixtures of ground-motion prediction equations as backbone models for a logic tree: an application to the subduction zone in Northern Chile, *Bulletin of Earthquake*

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Engineering, *13*(2), 483–501, http://doi.org/10.1007/s10518-014-9636-7.

Hanka, W., and R. Kind (1994), The GEOFON Program, http://doi.org/10.4401/ag-4196.

Hardebeck, J. L., and A. J. Michael (2006), Damped regional-scale stress inversions: Methodology and examples for southern California and the Coalinga aftershock sequence, *Journal of Geophysical Research: Solid Earth*, *111*(B11), n/a–n/a, http://doi.org/10.1029/2005JB004144.

Hauser, A. (1991), Hans Steffen, precursor del concepto Falla Liquiñe-Ofqui, *Andean Geology*, *18*(2), 177–179, http://doi.org/10.5027/andgeoV18n2-a07.

Heidbach, O., M. Tingay, A. Barth, J. Reinecker, D. Kurfeß, and B. Müller (2010), Global crustal stress pattern based on the World Stress Map database release 2008, *Tectonophysics*, *482*(1-4), 3–15, http://doi.org/10.1016/j.tecto.2009.07.023.

Ishibe, T., H. Tsuruoka, K. Satake, and M. Nakatani (2014), A focal mechanism solution catalog of earthquakes (M 2.0) in and around the Japanese Islands for 1985-1998, *Bulletin of the Seismological Society of America*, *104*(2), 1031–1036, http://doi.org/10.1785/0120130278.

Jaeger, C. (1979), *Rock mechanics and engineering*, Cambridge University Press.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1995), *Continuous Univariate Distributions, Vol. 2*, 2nd ed., Wiley.

Kagan, Y. Y. (1990), Random stress and earthquake statistics: spatial dependence, *Geophysical Journal International*, *102*(3), 573–583, http://doi.org/10.1111/j.1365-246X.1990.tb04584.x.

Kagan, Y. Y. (1991), 3-D rotation of double-couple earthquake sources, *Geophysical Journal International*, *106*(3), 709–716, http://doi.org/10.1111/j.1365-246X.1991.tb06343.x.

Kagan, Y. Y. (1992), Correlations of earthquake focal mechanisms, *Geophysical Journal International*, *110*(2), 305–320, http://doi.org/10.1111/j.1365-246X.1992.tb00876.x.

Kagan, Y. Y. (2007), Simplified algorithms for calculating double-couple rotation, *Geophysical Journal International*, *171*(1), 411–418, http://doi.org/10.1111/j.1365-246X.2007.03538.x.

Kagan, Y. Y. (2013), Double-couple earthquake source: symmetry and rotation, *Geophysical Journal International*, *194*(2), 1167–1179, http://doi.org/10.1093/gji/ggt156.

Kagan, Y. Y., and D. D. Jackson (2015), Likelihood analysis of earthquake focal mechanism distributions, *Geophysical Journal International*, *201*(3), 1409–1415, http://doi.org/10.1093/gji/ggv085.

Kanamori, H., and J. J. Cipar (1974), Focal process of the great Chilean earthquake May 22, 1960, *Physics of the Earth and Planetary Interiors*, *9*(2), 128–136, http://doi.org/10.1016/0031-9201(74)90029-6.

Kent, J. T. (1982), The Fisher-Bingham Distribution on the Sphere, *Journal of the Royal Statistical Society. Series B (Methodological)*, *44*(1), 71–80.

Kruskal, J. B. (1956), On the shortest spanning subtree of a graph and the traveling salesman problem, *Proceedings of the American Mathematical Society*, *7*(1), 48–48, http://doi.org/10.1090/S0002-9939-1956-0078686-7.

Kubo, H., W. Suzuki, S. Aoi, and H. Sekiguchi (2016), Source rupture process of the 2016 Kumamoto earthquake derived from strong motion data, *Japan Geoscience Union Meeting*, *2016*, MIS34–P62, http://doi.org/10.1186/s40623-016-0536-8.

Kushnir, A. F., V. M. Lapshin, V. I. Pinsky, and J. Fyen (1990), Statistically optimal event detection using small array data, *Bulletin of the Seismological Society of America*, *80*(6), 1934–1950.

Lin, J. (1991), Divergence measures based on the Shannon entropy, *IEEE Transactions on Information Theory*, *37*(1), 145–151, http://doi.org/10.1109/18.61115.

Lisle, Richard J. (2013), A critical look at the Wallace-Bott hypothesis in fault-slip analysis, *Bulletin de la Societe Geologique de France*, *184*(4-5), 299–306, http://doi.org/10.2113/gssgfbull.184.4-5.299.

Mardia, K. V., and P. E. Jupp (1999), *Directional Statistics*, 430 pp., John Wiley & Sons, Inc.

McKenzie, D. P. (1969), The relation between fault plane solutions for earthquakes and the directions of the principal stresses, *Bulletin of the Seismological Society of America*, *59*(2), 591–601.

Meier, M.-A., M. J. Werner, J. Woessner, and S. Wiemer (2014), A search for evidence of secondary static stress triggering during the 1992 M w 7.3 Landers, California, earthquake sequence, *Journal of Geophysical Research: Solid Earth*, *119*(4), 3354–3370, http://doi.org/10.1002/2013JB010385.

Michael, A. J. (1984), Determination of stress from slip data: Faults and folds, *Journal of Geophysical Research: Solid Earth*, *89*(B13), 11,517–11,526, http://doi.org/10.1029/JB089iB13p11517.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012

Michael, A. J. (1987), Use of focal mechanisms to determine stress: A control study, *Journal of Geophysical Research*, *92*(B1), 357, http://doi.org/10.1029/JB092iB01p00357.

Österreicher, F., and I. Vajda (2003), A new class of metric divergences on probability spaces and its applicability in statistics, *Annals of the Institute of Statistical Mathematics*, *55*(3), 639–653, http://doi.org/10.1007/BF02517812.

Park, B. U., and J. S. Marron (1990), Comparison of Data-Driven Bandwidth Selectors, *Journal of the American Statistical Association*, *85*(409), 66–72, http://doi.org/10.1080/01621459.1990.10475307.

Parzen, E. (1962), On Estimation of a Probability Density Function and Mode, *The Annals of Mathematical Statistics*, *33*(3), 1065–1076, http://doi.org/10.1214/aoms/1177704472.

Pascal, C. (2002), Interaction of faults and perturbation of slip: Influence of anisotropic stress states in the presence of fault friction and comparison between Wallace-Bott and 3D distinct element models, *Tectonophysics*, *356*(4), 307–322, http://doi.org/10.1016/S0040-1951(02)00413-4.

Pisarenko, V. F., A. F. Kushnir, and I. V. Savin (1987), Statistical adaptive algorithms for estimation of onset moments of seismic phases, *Physics of the Earth and Planetary Interiors*, *47*(C), 4–10, http://doi.org/10.1016/0031-9201(87)90062-8.

Pondrelli, S., S. Salimbeni, A. Morelli, G. Ekström, L. Postpischl, G. Vannucci, and E. Boschi (2011), European-Mediterranean Regional Centroid Moment Tensor catalog: Solutions for 2005-2008, *Physics of the Earth and Planetary Interiors*, *185*(3-4), 74–81, http://doi.org/10.1016/j.pepi.2011.01.007.

Rosenblatt, M. (1956), Remarks on Some Nonparametric Estimates of a Density Function, http://doi.org/10.1214/aoms/1177728190.

Sanderson, C., and R. Curtin (2016), Armadillo: a template-based C++ library for linear algebra, *The Journal of Open Source Software*, *1*(Sanderson 2010), 26, http://doi.org/10.21105/joss.00026.

Shannon, C. E. (1948), A Mathematical Theory of Communication, *Bell System Technical Journal*, *27*(3), 379–423, http://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Silver, P. G., and T. H. Jordan (1982), Optimal estimation of scalar seismic moment, *Geophysical Journal of the Royal Astronomical Society*, *70*(3), 755–787, http://doi.org/10.1111/j.1365-246X.1982.tb05982.x.

Stein, S., and M. E. Wysession (2003), *An Introduction to Seismology, Earthquakes, and Earth Structure*, 498 pp., Blackwell Publishing.

Strasser, F. O., M. C. Arango, and J. J. Bommer (2010), Scaling of the Source Dimensions of Interface and Intraslab Subduction-zone Earthquakes with Moment Magnitude, *Seismological Research Letters*, *81*(6), 941–950, http://doi.org/10.1785/gssrl.81.6.941.

Tucker, L. R. (1951), *A method for synthesis of factor analysis studies*, Washington, DC: Department of the Army.

von Neumann, J. (1951), *Various techniques used in connection with random digits*, vol. 12, 36–38 pp., U. S. Nat. Bureau of Standards.

Wallace, R. E. (1951), Geometry of Shearing Stress and Relation to Faulting, *The Journal of Geology*, *59*(2), 118–130, http://doi.org/10.1086/625831.

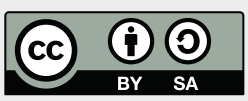
Warren, L. M. (2014), Dominant fault plane orientations of intermediate-depth earthquakes beneath South America, *Journal of Geophysical Research: Solid Earth*, *119*(7), 5762–5785, http://doi.org/10.1002/2013JB010856.

Wessel, P., W. H. F. Smith, R. Scharroo, J. Luis, and F. Wobbe (2013), Generic Mapping Tools: Improved version released, *EOS Trans. AGU*, *94*(45), 409–410, http://doi.org/10.1002/2013EO450001.

Zang, A., and O. Stephansson (2010), *Stress field of the Earth's crust*, Springer Science & Business Media.

Zoback, M. L. (1992), First- and second-order patterns of stress in the lithosphere: The World Stress Map Project, *Journal of Geophysical Research*, *97*(B8), 11,703, http://doi.org/10.1029/92JB00132.

STR 17/01. GFZ German Research Centre for Geosciences.
DOI: 10.2312/GFZb103-17012



ISSN 2190-7110