Geophysical Journal International

Geophys. J. Int. (2024) **237,** 1–13
GJI Seismology

https://doi.org/10.1093/gji/ggae011

Seismic moment tensor classification using elliptical distribution functions on the hypersphere

Mark J. Hoggard [1], Janice L. Scealy [2] and Brent G. Delbridge [3]

[1] *Research School of Earth Sciences, Australian National University, Canberra, ACT* 0200*, Australia. E-mail: mark.hoggard@anu.edu.au*
[2] *Research School of Finance, Actuarial Studies and Statistics, Australian National University, Canberra, ACT* 0200*, Australia*
[3] *EES-*17 *(Geophysics), Los Alamos National Laboratory, Los Alamos, NM* 87545*, USA*

Accepted 2024 January 7. Received 2023 December 20; in original form 2023 May 31

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

SUMMARY

Discrimination of underground explosions from naturally occurring earthquakes and other anthropogenic sources is one of the fundamental challenges of nuclear explosion monitoring. In an operational setting, the number of events that can be thoroughly investigated by analysts is limited by available resources. The capability to rapidly screen out events that can be robustly identified as not being explosions is, therefore, of great potential benefit. Nevertheless, possible mis-classification of explosions as earthquakes currently limits the use of screening methods for verification of test-ban treaties. Moment tensors provide a physics-based classification tool for the characterization of different seismic sources and have enabled the advent of new techniques for discriminating between earthquakes and explosions. Following normalization and projection of their six-degree vectors onto the hypersphere, existing screening approaches use spherically symmetric metrics to determine whether any new moment tensor may have been an explosion. Here, we show that populations of moment tensors for both earthquakes and explosions are anisotropically distributed on the hypersphere. Distributions possessing elliptical symmetry, such as the scaled von Mises–Fisher distribution, therefore provide a better description of these populations than the existing spherically symmetric models. We describe a method that uses these elliptical distributions in combination with a Bayesian classifier to achieve successful classification rates of 99 per cent for explosions and 98 per cent for earthquakes using existing catalogues of events from the western United States. The 1983 May 5 Crowdie underground nuclear test and 2018 July 20 DAG-1 deep-borehole chemical explosion are the only two explosions out of 140 that are incorrectly classified. Application of the method to the 2006–2017 nuclear tests in the Democratic People's Republic of Korea yields 100 per cent identification rates and we provide a simple routine `MTid` for general usage. The approach provides a means to rapidly assess the likelihood of an event being an explosion and can be built into monitoring workflows that rely on simultaneously assessing multiple different discrimination metrics.

Key words: Probability distributions; Statistical methods; Earthquake monitoring and test-ban treaty verification; Earthquake source observations; Theoretical seismology.

1 INTRODUCTION

Initial testing of nuclear weapons in the 1940s and their subsequent proliferation around the world has led to significant international effort at developing the capability to monitor nuclear testing. The majority of tests have been conducted underground (∼75 per cent) and are only evident from the seismic waves generated, which travel great distances and can be detected on seismometers. Global seismometer networks, such as the The International Monitoring System of the Comprehensive Nuclear-Test-Ban Treaty Organization, enable world-wide monitoring of underground nuclear explosions and provide crucial insights that cannot be easily estimated from other geophysical methods.

One of the principal challenges facing seismic monitoring concerns how to differentiate between explosions and earthquakes, the latter of which are far more common and can generate seismic energy with similar properties. Techniques that have been applied in this endeavour include analysis of event depth, the geometry of Rayleigh wave radiation patterns, the ratio of high-frequency *P*- to *S*-wave energy, comparison of local and coda-duration magnitudes

© The Author(s) 2024. Published by Oxford University Press on behalf of The Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

and, perhaps most extensively, comparison of body-wave (m_b) to surface-wave (M_S) magnitudes (e.g. Brune & Pomeroy 1963; Evernden *et al.* 1971; Blandford 1977; Walter *et al.* 1995; Koper *et al.* 2021; Carmichael 2021; Fisk *et al.* 2002). While these approaches have generally proven highly effective at identifying large underground explosions at teleseismic distances, no single method can successfully categorise all events (Pomeroy *et al.* 1982; Bowers & Selby 2009). For smaller events (i.e. $m_b < 4.5$), such as the 2006 and 2009 Democratic People's Republic of Korea (DPRK) nuclear tests, teleseismic M_s:m_b discrimination techniques have required modification (e.g. Selby *et al.* 2012). They have even failed for some large events, including the 2017 DPRK test, due to generation of anomalously large surface waves arising from interactions between the seismic source and local topography (Stevens & O'Brien 2018).

Spatial patterns of radiated seismic energy can be observed on seismometer networks and are generally distinct for explosions versus earthquakes. Explosions posses a significant isotropic element, in which compressional elastic energy radiates outwards in all directions from the source, whereas earthquakes are generally dominated by a source that radiates in four quadrants of alternating compressional and tensional elastic energy. Radiation patterns are therefore useful for inferring the seismic source, although all nuclear explosions contain some degree of non-isotropic radiation, complicating their utilization as simple seismic discriminants (Wallace 1991). Seismic sources can be described mathematically using the seismic moment tensor, a 3×3 symmetric tensor that describes the force couples operating at a point that would yield equivalent displacements to those of the real source (Gilbert 1970). The low-frequency seismic waves generated by earthquakes can be described by double-couple solutions that represent slip on a fault plane, whereas explosions and collapse events are described by cracks opening and/or closing in combination with isotropic volume changes (Julian *et al.* 1998). Recent development of efficient and robust methods to calculate the full seismic moment tensor has provided a new means to discriminate between different types of seismic event based on these characteristics (Dreger & Woods 2002; Minson & Dreger 2008; Ford *et al.* 2009).

Leveraging the ability of moment tensors to describe seismic sources, Ford *et al.* (2020) developed a novel statistical method for nuclear event 'screening', whereby any given event is screened out provided that a null hypothesis that the event was actually an explosion can be rejected. This method involves taking a catalogue of moment tensors for known explosions, individually normalizing each one and calculating its six-degree unit vector. This group of explosions thus defines a population of vectors on a 5-D manifold and Ford *et al.* (2020) show that a von Mises–Fisher probability distribution can be used to describe them, centred on their mean vector $\boldsymbol{\mu}$. The moment tensor for any new seismic event is also projected into this space and its angle from the mean vector, δ, is calculated. A core goal of event monitoring is that nuclear explosions are never missed, such that false classification of earthquakes as potential explosions is considered less problematic. Ford *et al.* (2020) showed that, for their data set of moment tensors for earthquakes and explosions, a cut-off angle of $\delta_M = 30.4^\circ$ was appropriate to ensure that any explosion would not be incorrectly screened out. Pasyanos & Chiang (2021) subsequently applied this screening method to a substantially larger catalogue of known earthquakes and explosions and showed that it was necessary to increase the cut-off angle to $\delta_M \approx 45^\circ$ to achieve optimal results.

In this study, we build on the pioneering work of Ford *et al.* (2020) to improve the robustness of this approach for event determination using seismic moment tensors. Specifically, we make two important changes. First, we demonstrate that the distribution of individual populations of moment tensor solutions are not spherically symmetric about their mean vector (i.e. they are anisotropically distributed). We show that a probability distribution function possessing elliptical symmetry, thereby accommodating different variances in each dimension, provides a better description of the population than the spherically symmetric distributions that have been previously used. Secondly, rather than providing a simple yes/no screening metric, we instead fit probability density functions to each of the known event populations (explosions and earthquakes) and assess the probability that any new event belongs to each one. This latter modification has the benefit of being easily adaptable, such that it is able to integrate other types of event population (e.g. collapse events), whilst also providing a probability measure that can be more readily built into existing probability-based workflows that rely on multiple independent event discriminants (e.g. Anderson *et al.* 2007).

2 EVENT DESCRIPTION ON THE HYPERSPHERE

2.1 Preliminaries and event catalogues

Much of the motivation and formalism of this study is taken directly from Ford *et al.* (2020, which in turn builds from Tape & Tape 2015) and we refer the reader to their paper for a complete description. Here, we will summarize only the basic essentials that are necessary to understanding the advances proposed in this study. We also make use of the same catalogues of known events that were employed in the two previous studies that investigated this approach. Moment tensors used by Ford *et al.* (2020) include 28 earthquakes from the western USA and 23 explosions that were assembled from previous compilations in the western USA (Ford *et al.* 2009) and DPRK (Chiang *et al.* 2018). The moment tensors used in Pasyanos & Chiang (2021) include 140 explosions (all from the Nevada National Security Site and listed in table 1 of that paper) and an inventory of ∼1200 earthquakes that occurred from 1992 to 2020 in the western United States, including Californian events from Long Valley (Dreger *et al.* 2000; Minson & Dreger 2008) and The Geysers geothermal field (Boyd *et al.* 2015), as well as across the broader region (Ford *et al.* 2009, and other sources). We are grateful to Michael Pasyanos and Doug Dreger for sharing the latest version of this earthquake catalogue, which, following removal of a small number of duplicates and incorrectly classified events that were included in their original study, now contains 1149 events. These two data sets are shown plotted on the fundamental lune of Tape & Tape (2012) in Fig. 1, alongside the canonical source mechanisms. We note that there appears to be a slightly greater number of earthquake solutions from the Pasyanos & Chiang (2021) data set extending into the compressional part of the lune than the dilational (i.e. the peripheral outliers are slightly biased upwards on Fig. 1b). We speculate that this behaviour might be attributed to several of these events coming from volcanic and hydrothermal fields in California, which are known to often possess a volumetric component in their moment tensor solution (e.g. Dreger *et al.* 2000; Boyd *et al.* 2015).

The 3×3 moment tensor is symmetric due to the requirement to conserve angular momentum and therefore has six unique elements on $\mathbb{R}^6$. Since we are interested here in determining source type rather than constraining source magnitude, it is helpful to normalize this matrix. One can consider unit vectors in p-dimensional Euclidean space, $\mathbb{R}^p$, as points on a $(p-1)$-dimensional hypersphere, $\mathbb{S}^{p-1}$, with unit radius and centre at the origin. The 5-D hypersphere,

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

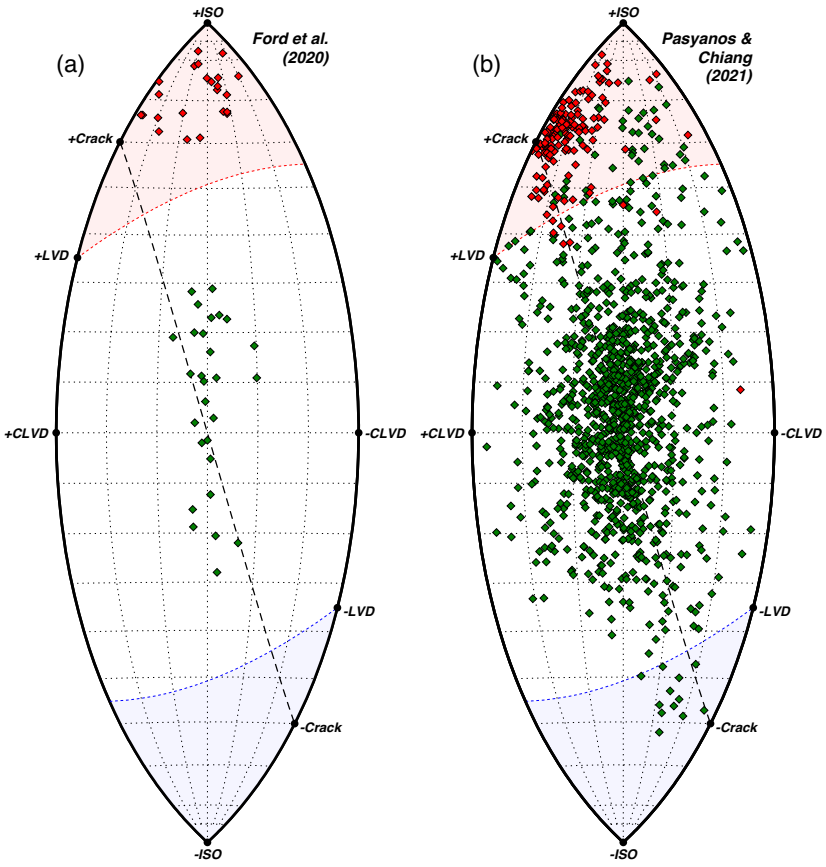

Figure 1. Seismic moment tensor catalogues projected onto the fundamental lune of Tape & Tape (2012). (a) Diamonds = data used in the event screening study of Ford *et al.* (2020); red = 23 explosions and green = 28 earthquakes. Black labelled circles correspond to canonical source mechanisms; +ISO/−ISO = isotropic compression/dilation; +Crack/−Crack = opening/closing crack with Poisson ratio of 0.25; +LVD/−LVD = positive/negative linear vector dipole; +CLVD/−CLVD = positive/negative compensated linear vector dipole; black dashed line = solutions that are a combination of crack plus double-couple tensors for Poisson ratio of 0.25; red/blue shaded area = regions where a classical beach-ball focal mechanism representation would be entirely compressional/dilational. (b) Same for the 140 explosions and 1149 earthquakes of Pasyanos & Chiang (2021). Lune projection is Hammer equal area.

$\mathbb{S}^5$, can therefore be mapped onto the space containing all 3 × 3 symmetric matrices with unit norm, $\hat{\mathbb{M}}$, via the norm-preserving isomorphism

$$m(\boldsymbol{x}) = \begin{pmatrix} x_1 & x_6/\sqrt{2} & x_5/\sqrt{2} \\ x_6/\sqrt{2} & x_2 & x_4/\sqrt{2} \\ x_5/\sqrt{2} & x_4/\sqrt{2} & x_3 \end{pmatrix} \tag{1}$$

of Silver & Jordan (1982), where $\boldsymbol{x} = (x_1, x_2, \ldots, x_6)^\top$ is the six-degree unit vector on $\mathbb{S}^5$.

2.2 Descriptive measures of event populations

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be points on $\mathbb{S}^5$ and consider the following general spherical location model

$$\begin{aligned} \boldsymbol{x}_i &= \boldsymbol{\Gamma}\boldsymbol{e}_i \quad \text{for} \quad i = 1, 2, \ldots, n \\ &= e_{1i}\boldsymbol{\mu} + (e_{2i}\boldsymbol{\gamma}_2 + e_{3i}\boldsymbol{\gamma}_3 + e_{4i}\boldsymbol{\gamma}_4 + e_{5i}\boldsymbol{\gamma}_5 + e_{6i}\boldsymbol{\gamma}_6) \end{aligned} \tag{2}$$

where n is the number of sampled events in the population and $\boldsymbol{\Gamma} = (\boldsymbol{\mu}, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \boldsymbol{\gamma}_4, \boldsymbol{\gamma}_5, \boldsymbol{\gamma}_6)$ is an orthogonal 6 × 6 matrix with first column equal to $\boldsymbol{\mu}$, which corresponds to the population mean direction. The vectors $\boldsymbol{e}_i = (e_{1i}, e_{2i}, e_{3i}, e_{4i}, e_{5i}, e_{6i})^\top \in \mathbb{S}^5$ are random errors on the hypersphere and have mean values for the population, $E(.)$, given by

$$\begin{aligned} E(e_{1i}) &= r > 0, \\ E(e_{2i}) = E(e_{3i}) = E(e_{4i}) &= E(e_{5i}) = E(e_{6i}) = 0, \end{aligned} \tag{3}$$

$$\begin{aligned} E(e_{ki}e_{ji}) &= 0 \quad \text{for} \quad j \neq k, \\ \text{and} \quad E(e_{2i}^2) \geq E(e_{3i}^2) &\geq E(e_{4i}^2) \geq E(e_{5i}^2) \geq E(e_{6i}^2). \end{aligned} \tag{4}$$

See Rivest (1984), Kent (1982) and Scealy *et al.* (2022) for further details. Under these assumptions, the true mean and dispersion matrix for $\boldsymbol{x}$ is

$$E(\boldsymbol{x}) = r\boldsymbol{\mu} \quad \text{and} \quad E\left(\boldsymbol{x}\boldsymbol{x}^\top\right) = \boldsymbol{\Gamma}\boldsymbol{D}\boldsymbol{\Gamma}^\top, \tag{5}$$

where $\boldsymbol{D}$ is a diagonal 6 × 6 matrix. The decomposition provided in the second line of eq. (2) is referred to as the tangent-normal decomposition. The vector $(e_{2i}\boldsymbol{\gamma}_2 + e_{3i}\boldsymbol{\gamma}_3 + e_{4i}\boldsymbol{\gamma}_4 + e_{5i}\boldsymbol{\gamma}_5 + e_{6i}\boldsymbol{\gamma}_6)$ is tangent to $\mathbb{S}^5$ at $\boldsymbol{\mu}$, while e_{1i} is equivalent to t in section 4.2 of Ford *et al.* (2020).

We can estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Gamma}$ using sample averages $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Gamma}}$, as in Kent (1982). The estimated mean direction is

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{i=1}^n \boldsymbol{x}_i}{\| \sum_{i=1}^n \boldsymbol{x}_i \|} \tag{6}$$

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

and the estimates of $\boldsymbol{\gamma}_2$, $\boldsymbol{\gamma}_3$, $\boldsymbol{\gamma}_4$, $\boldsymbol{\gamma}_5$, $\boldsymbol{\gamma}_6$ (denoted by $\hat{\boldsymbol{\gamma}}_2, \hat{\boldsymbol{\gamma}}_3, \hat{\boldsymbol{\gamma}}_4, \hat{\boldsymbol{\gamma}}_5, \hat{\boldsymbol{\gamma}}_6$, respectively) are the unit eigenvectors corresponding, in decreasing order, to the five strictly positive eigenvalues of the following sample dispersion matrix

$$\left(\boldsymbol{I}_6 - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\right)\left(\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i^\top\right)\left(\boldsymbol{I}_6 - \hat{\boldsymbol{\mu}}\hat{\boldsymbol{\mu}}^\top\right), \tag{7}$$

where $\boldsymbol{I}_6$ is the 6×6 identity matrix. Note that there is some non-uniqueness in the definition of $\hat{\boldsymbol{\Gamma}}$, such that any choice of the form $\{\hat{\boldsymbol{\mu}}, \pm\hat{\boldsymbol{\gamma}}_2, \pm\hat{\boldsymbol{\gamma}}_3, \ldots, \pm\hat{\boldsymbol{\gamma}}_6\}$ will suffice. If we wish to uniquely specify the signs of $\hat{\boldsymbol{\gamma}}_j$, it can be achieved with probability one by choosing, for example, the first component of each $\hat{\boldsymbol{\gamma}}_j$ to be positive. Parameters $\boldsymbol{\gamma}_2$, $\boldsymbol{\gamma}_3$, $\boldsymbol{\gamma}_4$, $\boldsymbol{\gamma}_5$ and $\boldsymbol{\gamma}_6$ have a convenient interpretation in that $\boldsymbol{\gamma}_2$ corresponds to the axis of highest variability in the tangent space at $\boldsymbol{\mu}$ (which is called the tangent plane to the surface at $\boldsymbol{\mu}$ for $\mathbb{S}^2$), $\boldsymbol{\gamma}_3$ is the next largest axis and so forth down to $\boldsymbol{\gamma}_6$, corresponding to the smallest variability direction. These parameters thus describe how the data are orientated on the surface of the hypersphere.

We can now estimate $\boldsymbol{\Gamma}$ for each of our individual data populations. The 23 explosion moment tensors from Ford *et al.* (2020) yield

$$\hat{\boldsymbol{\Gamma}} = \begin{pmatrix} 0.4507 & 0.4670 & 0.2222 & 0.3918 & 0.4747 & 0.3880 \\ 0.5234 & 0.5026 & -0.1368 & -0.1167 & -0.3934 & -0.5352 \\ 0.7135 & -0.6548 & 0.0667 & -0.2217 & 0.0635 & 0.0674 \\ 0.0271 & -0.2753 & -0.4267 & 0.7023 & 0.2274 & -0.4433 \\ 0.0243 & 0.1557 & -0.7597 & -0.4436 & 0.4227 & 0.1502 \\ -0.1124 & 0.0231 & 0.4101 & -0.3062 & 0.6209 & -0.5826 \end{pmatrix}$$

and the 140 explosion data in Pasyanos & Chiang (2021) yield

$$\hat{\boldsymbol{\Gamma}} = \begin{pmatrix} 0.3674 & 0.2891 & 0.0766 & 0.3604 & 0.2825 & 0.7523 \\ 0.4124 & 0.3783 & 0.0917 & 0.5701 & -0.1330 & -0.5794 \\ 0.8321 & -0.2855 & -0.1108 & -0.4583 & -0.0221 & -0.0575 \\ -0.0045 & -0.4356 & -0.7799 & 0.4472 & 0.0409 & 0.0195 \\ 0.0389 & -0.7080 & 0.6038 & 0.3635 & -0.0078 & 0.0204 \\ -0.0317 & -0.0267 & 0.0260 & -0.0543 & 0.9489 & -0.3071 \end{pmatrix}.$$

For earthquake moment tensors, the 28 events in Ford *et al.* (2020) give

$$\hat{\boldsymbol{\Gamma}} = \begin{pmatrix} -0.1253 & 0.3450 & 0.6231 & 0.5905 & 0.2621 & 0.2441 \\ -0.2547 & -0.4350 & -0.3882 & 0.5916 & -0.3151 & 0.3820 \\ 0.7456 & 0.4134 & -0.3500 & 0.1506 & -0.0272 & 0.3566 \\ -0.5946 & 0.6557 & -0.4414 & -0.1025 & 0.0435 & 0.0964 \\ -0.0263 & -0.2944 & -0.2362 & -0.0866 & 0.8816 & 0.2684 \\ 0.0962 & 0.0649 & -0.2964 & 0.5105 & 0.2283 & -0.7655 \end{pmatrix}$$

and the 1149 earthquakes from the database of Pasyanos & Chiang (2021) give

$$\hat{\boldsymbol{\Gamma}} = \begin{pmatrix} -0.4203 & 0.1344 & 0.8075 & 0.2949 & 0.2571 & 0.0090 \\ 0.8855 & 0.2226 & 0.3073 & 0.1830 & 0.1601 & -0.1132 \\ -0.0478 & -0.0129 & -0.3634 & 0.9230 & 0.0072 & 0.1162 \\ 0.1006 & 0.0435 & 0.0275 & -0.1075 & 0.1441 & 0.9772 \\ -0.0630 & -0.0048 & -0.2901 & -0.1078 & 0.9391 & -0.1354 \\ -0.1513 & 0.9645 & -0.1910 & -0.0667 & -0.0745 & -0.0183 \end{pmatrix}.$$

The first column of each $\hat{\boldsymbol{\Gamma}}$ is equivalent to the estimated mean direction of the population, $\hat{\boldsymbol{\mu}}$. We can immediately see that $\hat{\boldsymbol{\mu}}$ for the two explosion populations are similar to one another, but quite different to those for the earthquake data.

At this point, it is useful to explore the estimated errors (i.e. the deviation of each individual data point from the mean population vector along each of the principal axes), as they allow us to assess the validity of the model assumptions made at eqs (3) and (4). We can calculate

$$\hat{\boldsymbol{e}}_i = \hat{\boldsymbol{\Gamma}}^\top \boldsymbol{x}_i \quad \text{for} \quad i = 1, 2, \ldots, n \tag{8}$$

for each population of known events. We first explore simple histograms of $\hat{e}_{1i}$ in Fig. 2, before subsequently comparing each individual pair of the six components against one another. Fig. 3 shows the resulting relationship for the 23 explosions and 28 earthquakes used by Ford *et al.* (2020), while Fig. 4 is the equivalent plot for the 140 explosions and 1149 earthquakes from Pasyanos & Chiang (2021).

The first assumption that we need to check is whether the mean value of e_{1i} is greater than zero (i.e. $E(e_{1i}) > 0$). For a population of moment tensors drawn from a uniform distribution on $\mathbb{S}^5$, we would expect that the average of the $\hat{e}_{1i}$ values should be approximately equal to zero and this assumption would, therefore, not hold for that case. In the idealized case where the moment tensors come from a unimodal population on $\mathbb{S}^5$ with high-to-moderate concentration, we would expect $\hat{e}_{1i}$ to be left skewed and distributed relatively close to $+1$. As previously found by Ford *et al.* (2020), Fig. 2 shows that the explosion moment tensors are not uniformly distributed on the hypersphere. Whilst exhibiting a larger spread, the earthquake data show similar characteristics, with the distribution skewed towards a value of $+1$ in comparison to an equivalent synthetic data set that has been drawn from a uniform distribution on $\mathbb{S}^5$ (Fig. 2f). These results confirm that both the explosion and earthquake moment tensor populations appear to have $E(e_{1i}) > 0$, in support of the model assumption.

The other assumptions at eqs (3) and (4) are that the mean of the other errors (i.e. e_{2i}, e_{3i}, e_{4i}, e_{5i} and e_{6i}) should be zero and that these errors should be uncorrelated with one another. As a reminder, $\hat{e}_{2i}$ is, by definition, the component of highest variation in the space orthogonal to $(1, 0, 0, 0, 0, 0)^\top$ and each subsequent $\hat{e}$ should have correspondingly smaller variations. We first investigate whether these assumptions hold for the Ford *et al.* (2020) data set (Fig. 3). For the explosion data, we indeed see that the errors are fairly highly concentrated, roughly centred on zero, and that there is no obvious evidence of any correlation. The earthquake moment tensors are similarly centred on zero without correlations, but show substantially more dispersion (i.e. the data are less concentrated). For the larger event catalogues used in the study of Pasyanos & Chiang (2021), we see similar characteristics (Fig. 4).

Whilst the existing event catalogues therefore support the model assumptions, Figs 3 and 4 highlight an important additional property of these populations, which is that the variance of each $\hat{e}$ is not consistent for $\hat{e}_{2i}$ through $\hat{e}_{6i}$. This behaviour is particularly obvious in the explosion data, where the spread in $\hat{e}_{2i}$ and to some extent $\hat{e}_{3i}$ is generally larger than for the others. It is also visible in the earthquake data, for example in $\hat{e}_{2i}$ versus $\hat{e}_{6i}$. These panels are better described by elliptic contours of constant probability rather than circular ones—the implication of having different variances in orthogonal directions is that the data are anisotropically distributed.

In their study, Ford *et al.* (2020) propose using a von Mises–Fisher (vMF) distribution to describe these populations. The limitation of this description is that it is a spherically symmetric probability density function with a single variance parameter, κ, such that it has the same variance in each of the five dimensions. It is therefore best suited to describing isotropic data sets. To more accurately characterize anisotropic populations, we need a probability density function that allows each error direction to have its own variance.

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

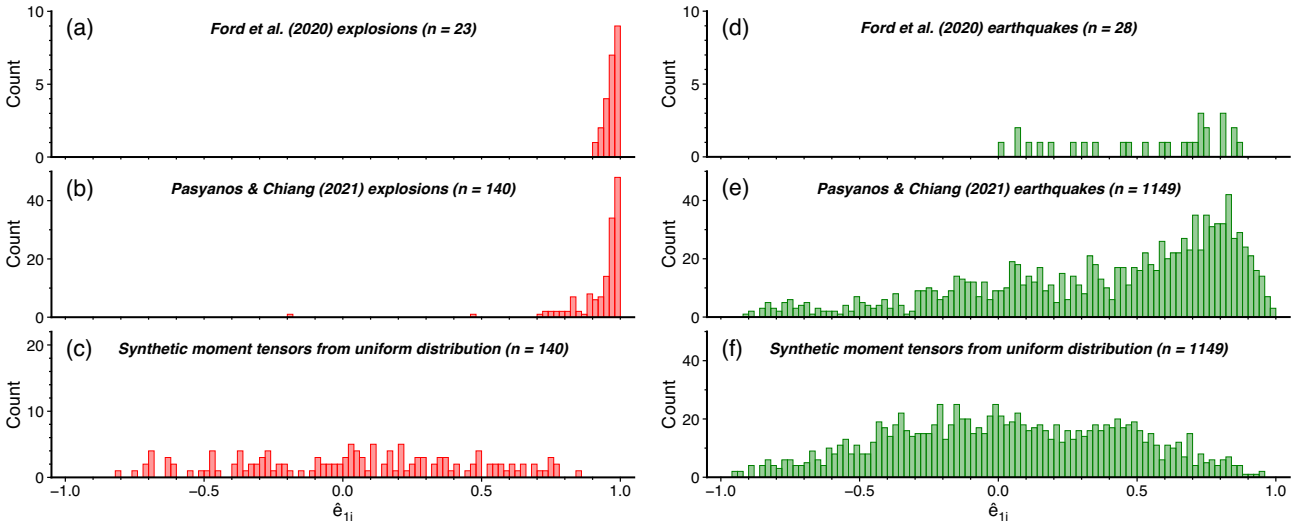

Figure 2. Histograms of $\hat{e}_{1i}$ for each of the event catalogues. (a) 23 explosions from Ford *et al.* (2020). (b) 140 explosions from Pasyanos & Chiang (2021). (c) 140 synthetic moment tensors generated by sampling a uniform distribution on $\mathbb{S}^5$. (d)–(f) Equivalent for earthquakes. Bin width in all histograms is 0.02.

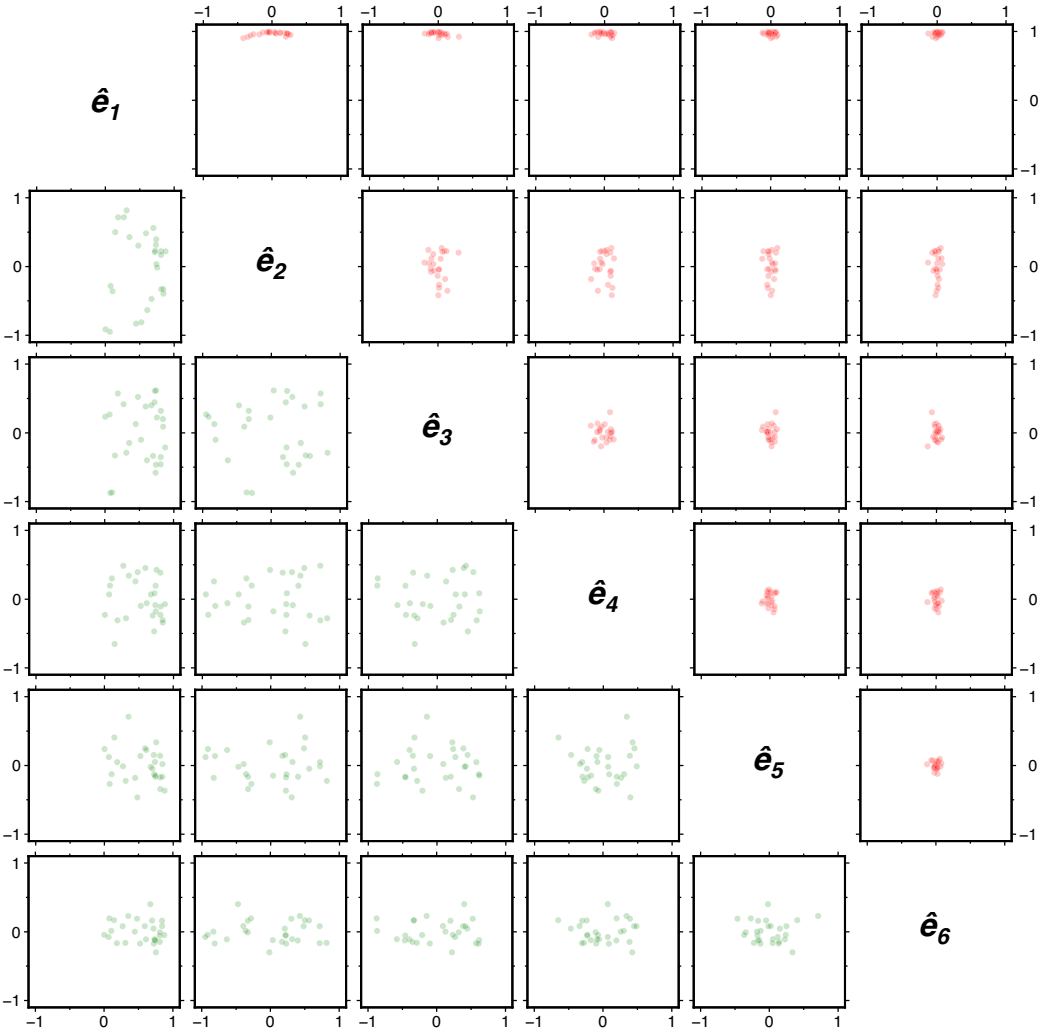

Figure 3. Relationship between each orthogonal pair of estimated errors, $\hat{e}_i$, for the Ford *et al.* (2020) event catalogues**.** Upper right (red) = 23 explosions; lower left (green) = 28 earthquakes; axes are consistent horizontally and vertically across each row and column; each individual data point is plotted with a constant level of transparency to highlight where results are more concentrated.

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

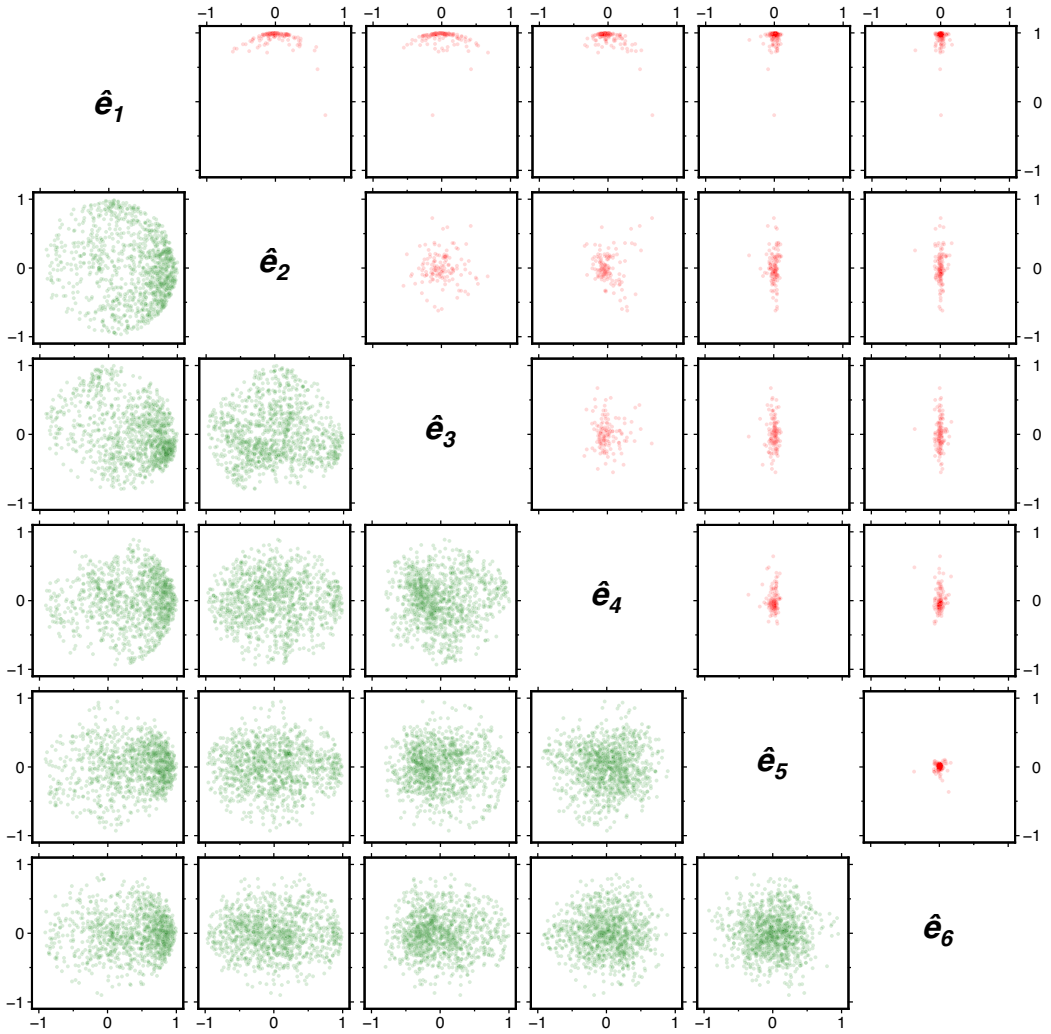

Figure 4. Relationship between each orthogonal pair of estimated errors, $\hat{\boldsymbol{e}}_i$, for the Pasyanos & Chiang (2021) event catalogues. Upper right (red) = 140 explosions; lower left (green) = 1149 earthquakes; axes are consistent horizontally and vertically across each row and column; each individual data point is plotted with a constant level of transparency to highlight where results are more concentrated.

2.3 Scaled von Mises–Fisher distribution

Scealy & Wood (2019) have proposed a new distribution to characterize data populations on $\mathbb{S}^{p-1}$ for dimensions, $p \geq 3$, which is called the scaled von Mises–Fisher (SvMF) distribution. It possesses ellipse-like symmetry and contains the same number of parameters as the $(p-1)$-dimensional normal distribution. It also contains a normalizing factor that ensures the probability density function integrates to one over the hypersphere and is thus a valid probability density. Importantly, this normaliation factor is easy to compute, unlike for Kent distributions (e.g. Paine *et al.* 2018), making estimation of the parameters computationally straightforward. Simulation is also relatively easy because it simply requires rescaling of a regular vMF random variable.

A unit random vector $\boldsymbol{x}$ has the $(p-1)$-dimensional SvMF distribution if its probability density function is

$$f(\boldsymbol{x};\boldsymbol{\Gamma};\boldsymbol{a},\kappa) = \left\{\frac{\kappa^{(p/2)-1}}{(2\pi)^{p/2} I_{(p/2)-1}(\kappa)}\right\}\left\{(\boldsymbol{x}^\top\boldsymbol{\mu})^2 + \sum_{j=2}^{p}\left(\frac{\boldsymbol{x}^\top\boldsymbol{\gamma}_j}{a_j}\right)^2\right\}^{-\frac{p-1}{2}} \times \exp\left\{\frac{\kappa\boldsymbol{x}^\top\boldsymbol{\mu}}{\left[(\boldsymbol{x}^\top\boldsymbol{\mu})^2 + \sum_{j=2}^{p}\left(\frac{\boldsymbol{x}^\top\boldsymbol{\gamma}_j}{a_j}\right)^2\right]^{1/2}}\right\}, \quad (9)$$

where $\boldsymbol{\Gamma} = (\boldsymbol{\mu}, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \ldots, \boldsymbol{\gamma}_p)$ are location parameters with the same definitions used in Section 2.2, I_ν denotes the modified Bessel function of the first kind and order ν, κ is a concentration parameter and $\boldsymbol{a} = (a_2, a_3, \ldots, a_p)^\top$ are shape parameters that satisfy

$$\prod_{j=2}^{p} a_j = 1, \quad a_j > 0, \quad \text{for} \quad j = 2, 3, \ldots, p. \quad (10)$$

Note that if $a_j = 1$ for all $j = 2, 3, \ldots, p$, then $(\boldsymbol{x}^\top\boldsymbol{\mu})^2 + \sum_{j=2}^{p}\left(\boldsymbol{x}^\top\boldsymbol{\gamma}_j/a_j\right)^2 = 1$ and the density function of eq. (9) becomes equivalent to that of the regular vMF probability density function. The SvMF density is therefore essentially a rescaled vMF density that allows for anisotropy (i.e. elliptical contours) via the shape parameter vector, $\boldsymbol{a}$.

2.4 Density function parameter estimation

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ be a random sample from the SvMF distribution with probability density function given by eq. (9) with $p = 6$. The location parameter matrix $\boldsymbol{\Gamma}$ can be estimated using the sample mean and dispersion matrix, as previously outlined in Section 2.2, which are called moment estimators. The concentration parameter κ

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

and the shape parameter vector $\boldsymbol{a}$ can be estimated using maximum-likelihood estimation, as summarized below (see Scealy & Wood 2019, for further details).

We begin by defining the 6×6 orthogonal matrix

$$\boldsymbol{H}(\boldsymbol{\mu}) = \begin{pmatrix} \mu_1 & \boldsymbol{\mu}_L^\top \\ \boldsymbol{\mu}_L & \frac{1}{1+\mu_1}\boldsymbol{\mu}_L\boldsymbol{\mu}_L^\top - \boldsymbol{I}_5 \end{pmatrix} = \{\boldsymbol{\mu}, \boldsymbol{H}^*(\boldsymbol{\mu})\}, \tag{11}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6)^\top$, $\boldsymbol{\mu}_L = (\mu_2, \mu_3, \mu_4, \mu_5, \mu_6)^\top$ and $\boldsymbol{H}^*(\boldsymbol{\mu})$ is a 6×5 matrix whose columns are orthogonal to $\boldsymbol{\mu}$. Let $\boldsymbol{K}^*$ be a general 5×5 orthogonal matrix defined such that

$$\boldsymbol{\Gamma} = \{\boldsymbol{\mu}, \boldsymbol{H}^*(\boldsymbol{\mu})\boldsymbol{K}^*\} \tag{12}$$

holds. Then, let

$$\boldsymbol{V} = \boldsymbol{K}^* \mathrm{Diag}\left[\left(a_2^2, a_3^2, a_4^2, a_5^2, a_6^2\right)^\top\right] \boldsymbol{K}^{*\top}, \tag{13}$$

where $\mathrm{Diag}(\boldsymbol{v})$ is a diagonal matrix with vector $\boldsymbol{v}$ on the diagonal, $\boldsymbol{V}$ is a 5×5 dimensional symmetric positive definite matrix with constraint that $\det(\boldsymbol{V}) = 1$, which corresponds to the condition in eq. (10), and assume $a_2 > a_3 > a_4 > a_5 > a_6$.

If $\boldsymbol{\mu}$ and κ are known, then the log-likelihood for $\boldsymbol{V}$ is

$$-\frac{5}{2}\sum_{i=1}^{n} \log\left(e_{1i}^2 + \boldsymbol{e}_i^{*\top}\boldsymbol{V}^{-1}\boldsymbol{e}_i^*\right) + \sum_{i=1}^{n} \frac{\kappa e_{1i}}{\left(e_{1i}^2 + \boldsymbol{e}_i^{*\top}\boldsymbol{V}^{-1}\boldsymbol{e}_i^*\right)^{\frac{1}{2}}}, \tag{14}$$

where $e_{1i} = \boldsymbol{\mu}^\top \boldsymbol{x}_i$ and $\boldsymbol{e}_i^* = \boldsymbol{H}^*(\boldsymbol{\mu})^\top \boldsymbol{x}_i$. To estimate $\boldsymbol{V}$, we maximize eq. (14) with respect to $\boldsymbol{V}$ subject to the constraint that $\det(\boldsymbol{V}) = 1$ or equivalently $\log\{\det(\boldsymbol{V})\} = 0$. For a suitable starting value (such as $\hat{\boldsymbol{V}}_0 = \boldsymbol{I}_5$), this approach leads to the iterative reweighting algorithm

$$\hat{\boldsymbol{V}}_{m+1} = \frac{\sum_{i=1}^{n}\left[5\left(\hat{s}_{i(m)}^2\right)^{-1} + \kappa e_{1i}\left(\hat{s}_{i(m)}^2\right)^{-\frac{3}{2}}\right]\boldsymbol{e}_i^*\boldsymbol{e}_i^{*\top}}{\left\{\det\left(\sum_{i=1}^{n}\left[5\left(\hat{s}_{i(m)}^2\right)^{-1} + \kappa e_{1i}\left(\hat{s}_{i(m)}^2\right)^{-\frac{3}{2}}\right]\boldsymbol{e}_i^*\boldsymbol{e}_i^{*\top}\right)\right\}^{\frac{1}{5}}}, \tag{15}$$

where m is the iteration number and $\hat{s}_{i(m)}^2 = e_{1i}^2 + \boldsymbol{e}_i^{*\top}\hat{\boldsymbol{V}}_m^{-1}\boldsymbol{e}_i^*$. Note that eq. (15) needs to be iterated until convergence, conditional on $\boldsymbol{\mu}$ and κ remaining fixed at their current values during each iteration loop. Updating eq. (15) is trivial, since it simply involves calculating weighted sums-of-squares where the weights are updated at each iteration.

If $\boldsymbol{\mu}$ and $\boldsymbol{V}$ are known, then the log-likelihood for κ is

$$-n\log\left(\frac{(2\pi)^3 I_2(\kappa)}{\kappa^2}\right) + \kappa\sum_{i=1}^{n}\boldsymbol{x}_i^\top\boldsymbol{\mu}\left[\left(\boldsymbol{x}_i^\top\boldsymbol{\mu}\right)^2 + \boldsymbol{x}_i^\top\boldsymbol{H}^*(\boldsymbol{\mu})\boldsymbol{V}^{-1}\boldsymbol{H}^*(\boldsymbol{\mu})^\top\boldsymbol{x}_i\right]^{-\frac{1}{2}}. \tag{16}$$

The modified Bessel function of the first kind is available in many software packages, rendering the above log-likelihood straightforward to maximize with respect to κ by applying one-dimensional, derivative-free interval search methods. Given $\boldsymbol{\mu}$ for a population, in order to compute joint estimates of κ and $\boldsymbol{V}$, we iterate between maximizing eqs (14) and (16), where the most recent update of $\boldsymbol{V}$ is used in eq. (16) and the most recent update of κ is used in eq. (14), until both sets of parameters converge. Once the final estimate of $\boldsymbol{V}$ is reached, we extract its eigenvalues in order of largest to smallest, giving $\hat{a}_2^2 > \hat{a}_3^2 > \hat{a}_4^2 > \hat{a}_5^2 > \hat{a}_6^2$, and take their square roots to obtain estimates of a_2, a_3, a_4, a_5 and a_6.

We have applied this procedure to the individual event catalogues, starting with the 23 explosions in Ford *et al.* (2020), yielding

$$\hat{\kappa} = 124.9778, \quad \hat{\boldsymbol{a}} = (2.3447, 1.2361, 1.0865, 0.6350, 0.5001)^\top$$

and also for the 140 explosions in Pasyanos & Chiang (2021), which give

$$\hat{\kappa} = 54.6537, \quad \hat{\boldsymbol{a}} = (2.0725, 2.0527, 1.2762, 0.4786, 0.3849)^\top.$$

Equivalent values for the 28 earthquake moment tensors in Ford *et al.* (2020) are

$$\hat{\kappa} = 4.6659, \quad \hat{\boldsymbol{a}} = (1.9875, 1.4970, 0.9693, 0.7320, 0.4737)^\top$$

and the 1149 earthquakes of Pasyanos & Chiang (2021) give

$$\hat{\kappa} = 2.6088, \quad \hat{\boldsymbol{a}} = (1.5258, 1.0745, 0.9805, 0.8035, 0.7742)^\top.$$

These estimates are compatible with our earlier visual analysis in Section 2.2. The explosions have a higher concentration parameter, κ, estimate in comparison to the more dispersed earthquakes. The estimate of $\boldsymbol{a}$ for explosions is far from a value of $(1, 1, 1, 1, 1)^\top$, indicating that the distribution possesses reasonably high levels of anisotropy. The two earthquake populations have $\hat{\boldsymbol{a}}$ slightly closer to $(1, 1, 1, 1, 1)^\top$. Based off the sizes of each component, the smaller data set from Ford *et al.* (2020) displays anisotropy that is not dissimilar to the explosions, whereas the larger one from Pasyanos & Chiang (2021) possesses more moderate levels of anisotropy.

2.5 Simulation of synthetic populations

We can draw synthetic populations from an SvMF probability distribution (hereafter simply referred to as the 'elliptical' distribution) using the following three-step algorithm to generate the random variable $\boldsymbol{x}$ on $\mathbb{S}^5$:

(i) Generate $\boldsymbol{z} = (z_1, z_2, z_3, z_4, z_5, z_6)^\top \in \mathbb{S}^5$ from the vMF distribution (hereafter 'isotropic' distribution) with concentration parameter κ and mean direction $\boldsymbol{\mu} = (1, 0, 0, 0, 0, 0)^\top$. Algorithm 1 in the appendix of Ford *et al.* (2020) can be used here, which is taken from Wood (1994).

(ii) Let $\boldsymbol{e} = \left(z_1^2 + \sum_{i=2}^{6} z_i^2 a_i^2\right)^{-\frac{1}{2}} (z_1, z_2 a_2, z_3 a_3, z_4 a_4, z_5 a_5, z_6 a_6)^\top$.

(iii) Now let $\boldsymbol{x} = \boldsymbol{\Gamma}\boldsymbol{e}$.

In Fig. 5, we demonstrate this procedure by generating synthetic populations of explosions and earthquakes from the fitted probability distributions of both the isotropic and elliptical models applied to the Pasyanos & Chiang (2021) data sets. We randomly draw 10 000 events from each of the explosion and earthquake distributions and project them onto the fundamental lune. Comparison of these two population descriptors with the real moment tensor solutions visually demonstrates the improvement gained by using an elliptical probability density function that can accommodate anisotropy within the population distribution. For the explosion data in particular, a population drawn from the elliptical model exhibits characteristics on the fundamental lune that are more similar to the true data than that for the isotropic model. The impact on the distribution of earthquakes is more modest and reflects the lower level of anisotropy in accordance with the inferred value of $\hat{\boldsymbol{a}}$. We can also see that the synthetic earthquakes are more diffuse than the real data, indicating that the distribution of moment tensor solutions is, in reality, more 'heavy-tailed' than suggested by the SvMF model. Scealy & Wood (2019) have demonstrated that introduction of an additional shape parameter (their a_1) can suitably modify the SvMF to accommodate this property. Estimating its value is, however, non-trivial and also has limited impact on the performance of our event classification scheme. We have therefore elected to use the standard SvMF for our elliptical model in this study.

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

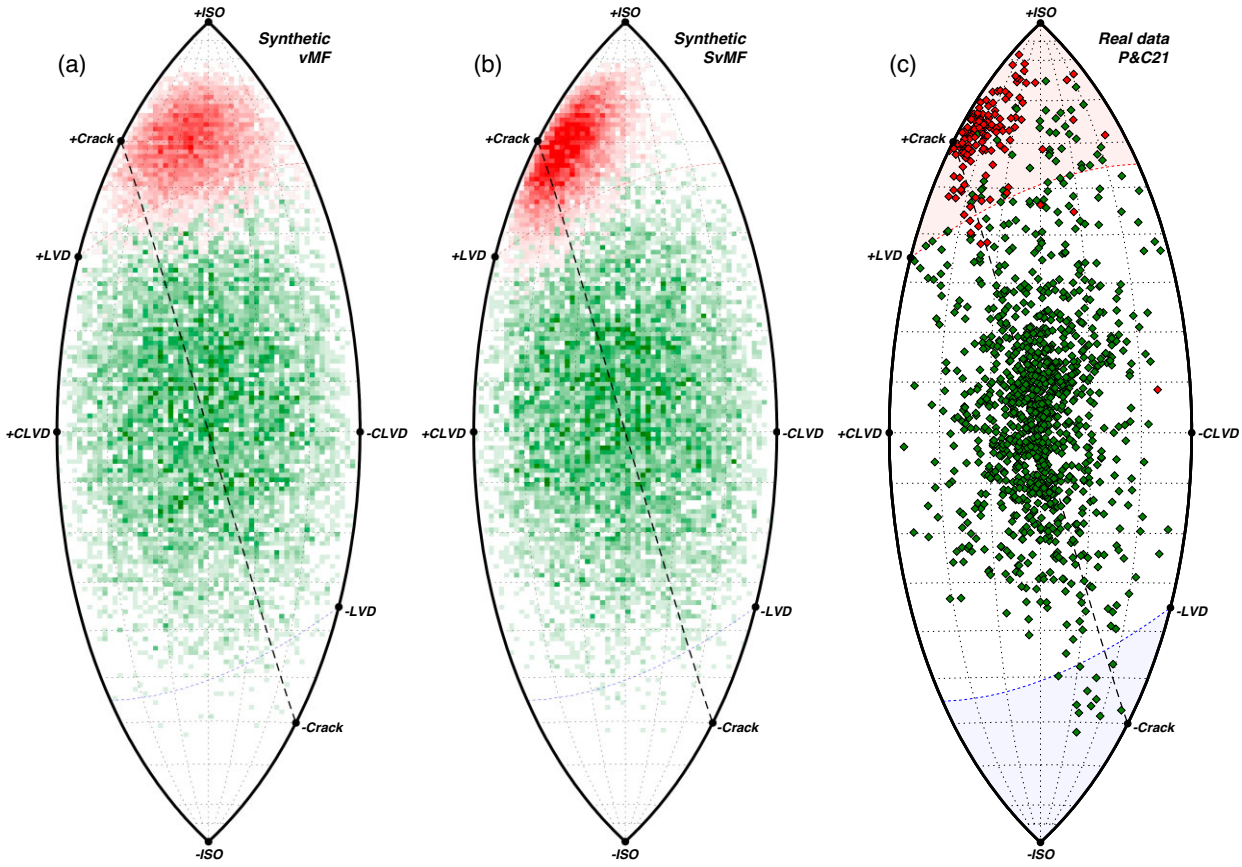

Figure 5. Comparison of synthetic populations generated from probability density functions with real moment tensor solutions. (a) Count maps for 10 000 explosions and 10 000 earthquakes randomly drawn from vMF distributions (isotropic) that have been fitted to the event catalogues of Pasyanos & Chiang (2021). (b) Same for the SvMF distribution (elliptical). (c) Real moment tensor solutions from Pasyanos & Chiang (2021). All data have been projected onto the fundamental lune of Tape & Tape (2012), which is plotted using a Hammer equal area projection.

Finally, adopting a single vMF or SvMF model to represent each population works optimally when the data are unimodally distributed on the hypersphere with approximately symmetric errors in each of the other axes (i.e. e_{2i}, e_{3i}, e_{4i}, e_{5i} and e_{6i} are not skewed). While we do not find evidence for bimodality in the population distributions (Figs 3 and 4), there is some evidence for skew (e.g. $\hat{e_3}$ values for the earthquakes in Fig. 4 are slightly right-skewed). These properties indicate that introducing additional parameters that would further influence the shape of the function, or even potentially fitting a mixture of functions, may improve model fits to the earthquake population distribution. Nevertheless, they would render estimation of the various model parameters substantially more complex. Given that the most important aspect for improving classification rates is recognizing that these populations are anisotropically distributed, we therefore confine ourselves here to simply exploring the benefits of introducing a function with elliptical symmetry rather than the isotropic one that has been previously used.

3 EVENT DISCRIMINATION

In terms of the formulation presented here, the task of event discrimination requires discerning which of the known populations any new event, $\boldsymbol{x}$, likely belongs to. Moreover, in order to combine our assessment with those of other potential discrimination techniques (e.g. Anderson *et al.* 2007), it is helpful to provide a level-of-certainty or probability measure alongside this assessment. As in Ford *et al.* (2020), our procedure uses training catalogues of known events to constrain population properties and is therefore an example of classification/supervised learning/discrimination (see section 4.3 of Hastie *et al.* 2009, for further details).

Let H be the number of different types of event for which we have a training population and $\hat{\boldsymbol{\Gamma}}_g$, $\hat{\kappa}_g$ and $\hat{\boldsymbol{a}}_g$ be the estimates of the respective parameters $\boldsymbol{\Gamma}$, κ and $\boldsymbol{a}$ for the elliptical model fitted to each population $g = 1, 2, \ldots, H$. Note that throughout this study, we have used two event types (i.e. $H = 2$, with $g = 1$ for explosions and $g = 2$ for earthquakes), but this approach can be generalized to include other event types (e.g. collapse events). We will again use the aforementioned event catalogues of Pasyanos & Chiang (2021) as our training populations in this analysis.

The conditional probability density of $\boldsymbol{x}$ belonging to any group is $f(\boldsymbol{x}; \hat{\boldsymbol{\Gamma}}_g, \hat{\boldsymbol{a}}_g, \hat{\kappa}_g)$ and π_g is the prior probability of it being in that group, where $\sum_{g=1}^{H} \pi_g = 1$. A simple application of Bayes theorem gives

$$P(G = g|\boldsymbol{x}) \quad = \quad \frac{\pi_g f(\boldsymbol{x}; \hat{\boldsymbol{\Gamma}}_g, \hat{\boldsymbol{a}}_g, \hat{\kappa}_g)}{\sum_{g=1}^{H} \pi_g f(\boldsymbol{x}; \hat{\boldsymbol{\Gamma}}_g, \hat{\boldsymbol{a}}_g, \hat{\kappa}_g)}, \tag{17}$$

where $P(G = g|\boldsymbol{x})$ is the probability that an event $\boldsymbol{x}$ belongs to the group $G = g$ and G is the random variable representing the groups.

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

Table 1. Comparison of classification success rates for 140 explosions and 1149 earthquakes from the Pasyanos & Chiang (2021) event catalogues using different screening methods.

Screening approach	Earthquakes (per cent)	Explosions (per cent)
$\delta_M = 30.4^\circ$	98	82
$\delta_M = 45^\circ$	94	97
vMF ($\pi_1 = 0.5$)	97	95
SvMF ($\pi_1 = 0.5$)	99	96
SvMF ($\pi_1 = 0.95$)	99	97
SvMF ($\pi_1 = 0.9999$)	98	99

Notes. δ_M refers to screening using a spherically symmetric cut-off angle (the Ford *et al.* 2020, approach); vMF refers to our revised highest-likelihood approach applied to spherically isotropic von Mises–Fisher distributions; SvMF refers to our revised highest-likelihood approach applied to elliptical scaled von Mises–Fisher distributions; π_1 is the prior probability that each event was an explosion.

3.1 Event screening with uniform priors

As explained by Ford *et al.* (2020), event screening is a more nuanced form of explosion monitoring practice that aims to identify whether a new event belongs to the population of explosions ('screened in') or not ('screened out'). We can see from eq. (17) that, for application to our case, an event can be screened in as an explosion if

$$\pi_1 f(\boldsymbol{x}; \hat{\boldsymbol{\Gamma}}_1; \hat{\boldsymbol{a}}_1, \hat{\kappa}_1) > \pi_{g\neq 1} f(\boldsymbol{x}; \hat{\boldsymbol{\Gamma}}_{g\neq 1}; \hat{\boldsymbol{a}}_{g\neq 1}, \hat{\kappa}_{g\neq 1}) \quad \text{for} \quad g = 2, 3, \ldots, H. \tag{18}$$

In other words, it is more likely to belong to group 1 (explosions) than any other group. In the first instance, we will assume equal priors (i.e. $\pi_1 = \pi_2 = \frac{1}{2}$), such that our classifier will just be comparing likelihoods. In the following section, we will explore the impact of making other choices.

Our screening approach based on eq. (18) differs from that adopted by Ford *et al.* (2020), who also define a classifier, but only train it on the explosion data (i.e. it does not involve modelling the earthquake distribution). Their screening rule is instead based on calculating the angle between a new event $\boldsymbol{x}$ and the mean direction of the explosion data according to

$$\delta = \arccos\left(\boldsymbol{x}^\top \hat{\boldsymbol{\mu}}\right), \tag{19}$$

where they use $\hat{\boldsymbol{\mu}} = (0.450, 0.524, 0.713, 0.272, 0.0245, -0.112)^\top$, which is virtually identical to our result for their explosion data set in Section 2.2. If δ is either smaller or larger than a cut-off value of $\delta_M = 30.4^\circ$, the new event is either screened in or out, respectively. This cut-off angle was selected based on that angular distance containing 99.9 per cent of their cumulative distribution function (i.e. a spherically symmetric vMF distribution) for $\hat{e}_1$ (t in their nomenclature). When Pasyanos & Chiang (2021) applied it to their significantly larger catalogues of events, they suggested δ_M should be increased to at least 45°, which yielded misidentification rates of ∼3 per cent for explosions and ∼6 per cent for earthquakes. The reason for needing to make this adjustment is evident in Fig. 1, where the populations of explosions and earthquakes are well separated for the Ford *et al.* (2020) catalogues, but exhibit substantially more overlap for Pasyanos & Chiang (2021).

In Table 1, we list mis-classification rates for the training event catalogues of Pasyanos & Chiang (2021) based on the Ford *et al.* (2020) screening method and our revised approach. Note that, in order to separate out the effects of adopting maximum likelihood as a screening criterion from those associated with using an elliptical rather than isotropic function, we have also tested using maximum likelihood with a vMF distribution (entry 'vMF $\pi_1 = 0.5$' in Table 1). This modification requires replacing the SvMF densities with the fitted vMF densities in eq. (17) for the two training groups. Results for individual moment tensors, with false positives and false negatives highlighted, are shown in Fig. 6.

Note that our mis-classification rates for the $\delta_M = 30.4^\circ$ and 45° cut-off methods are in close agreement with rates found by Pasyanos & Chiang (2021). The former does a good job of classifying earthquakes (98 per cent correctly assigned), but misses 18 per cent of the explosions. The latter, on the other hand, is a better choice for classifying explosions (97 per cent correctly characterized), but incorrectly identifies 6 per cent of earthquakes (67 out of 1149) as belonging to the explosion population. Whilst the latter is therefore better suited to the overarching goal of nuclear event monitoring (i.e. to not miss any potential tests), it comes with the associated cost of having to investigate many potential explosions that turn out to be earthquakes. It is apparent that, for the isotropic distribution, switching the classifier from one based on a cut-off angle to a maximum-likelihood (eq. 18) returns results that are intermediate to these two cases, but closer to the optimal result for each. Seven of 140 explosions are missed (compared to four for the $\delta_M = 45^\circ$ cut-off), while only 32 of 1149 earthquakes are misidentified as explosions (compared to 23 for the $\delta_M = 30.4^\circ$ cut-off).

Comparing screening results that use the isotropic distribution to those that use the elliptical version reveals an improvement in classification results for both types of event. For the elliptical case, 1137 of 1149 earthquakes are correctly assigned, whilst only 5 of 140 explosions are missed. Interestingly, of those missed explosions, one returns a probability of belonging to the explosion group of 38.8 per cent, which is not far from the 50 per cent threshold (N.B., the other four return probabilities of less than 0.1 per cent). In comparison, for the 12 mis-classified earthquakes, three return probabilities of being earthquakes that are greater than 10 per cent (42.0, 34.1 and 15.2 per cent). These values are significant and would likely raise attention because not one of the 134 correctly assigned explosions returned a probability of less than 80 per cent (the lowest is 81.0 per cent and second lowest is 94.6 per cent). These outlying events underline the potential benefit of considering probability values over a simplified yes/no screening statistic—closer investigation of the small group that straddle the boundary between populations highlights approximately one-quarter of all events that were mis-classified for these catalogues.

3.2 Tweaking the priors

We next explore the effect of modifying the prior probability that an event was an explosion (i.e. π_1). In terms of seismic event monitoring, the balance being struck is between the number of real explosions missed versus the number of earthquakes that are incorrectly classified as explosions, and different users are likely to have different requirements regarding this trade-off. The classifier proposed in eq. (18) lends itself naturally to this issue because it can be easily adjusted.

We demonstrate this aspect using three different values of $\pi_1 =$ 0.5, 0.95 and 0.9999, which progressively increase the prior probability that an event belongs to the explosion population (Table 1 and Fig. 7). As a reminder, the benchmark elliptical case with equal priors mis-classified 5 of 140 explosions and 12 of 1149 earthquakes. Setting $\pi_1 = 0.95$ changes these numbers to 4 and 14, respectively, while setting $\pi_1 = 0.9999$ yields 2 and 21 (equivalent to successful classification rates of 98.6 per cent for explosions and 98.2 per cent for earthquakes). Thus, it is clear that adjusting π_1 provides an intuitive and simple way to alter the balance between reducing the

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

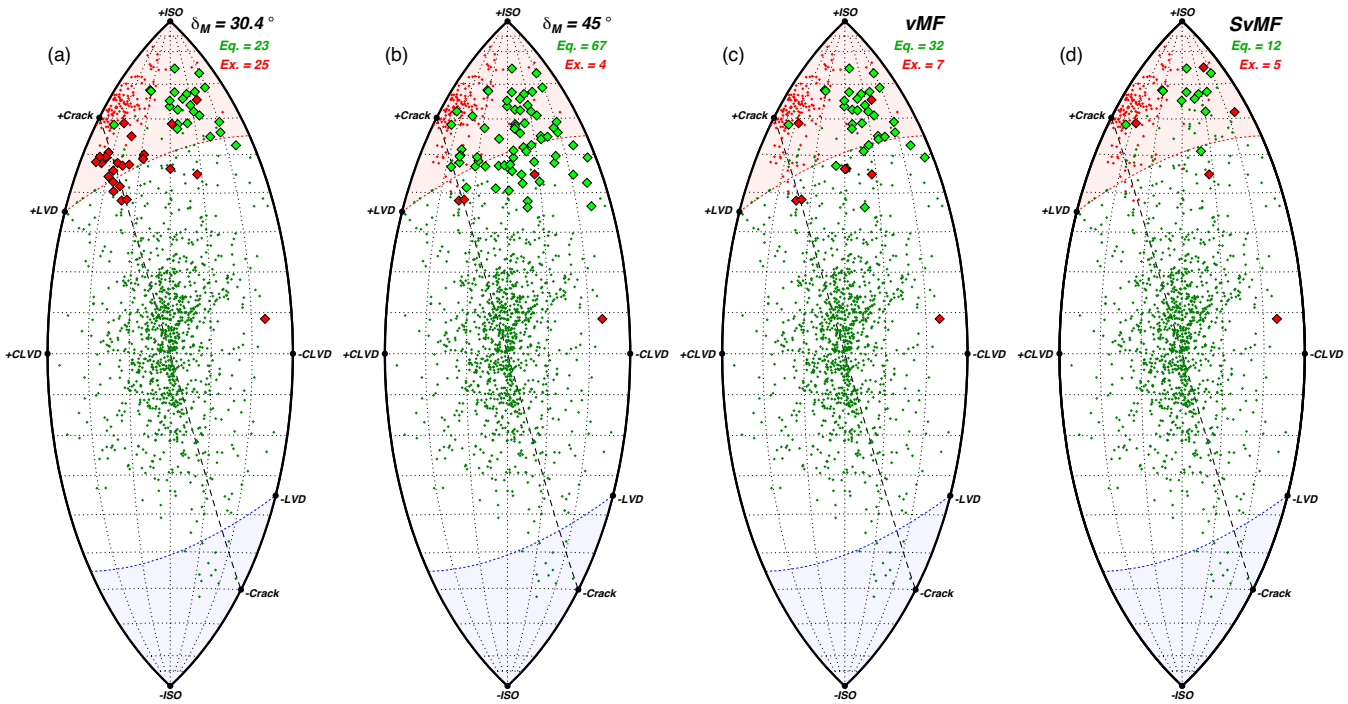

Figure 6. Mis-classification of event source type for different screening methods. (a) 140 explosions and 1149 earthquakes from Fig. 1(b) (Pasyanos & Chiang 2021) screened using a spherically symmetric cut-off angle of $\delta_M = 30.4°$ (i.e. the original Ford *et al.* 2020 scheme). Large red diamonds = explosions (Ex.) that are incorrectly classified as earthquakes; large green diamonds = earthquakes (Eq.) incorrectly classified as explosions; inset lists the number of mis-classified events. (b) Same for a cut-off angle of $\delta_M = 45°$. (c) Same when spherically symmetric vMF distributions are individually fitted to the explosion and earthquake populations and the cut-off angle classifier is replaced with a maximum-likelihood approach according to eq. (18). (d) Same for the SvMF distribution. All data have been projected onto the fundamental lune of Tape & Tape (2012), which is plotted using a Hammer equal area projection.

number of explosions missed versus the number of earthquakes incorrectly classified as explosions. Comparison of the latter case against the original screening approach of Ford *et al.* (2020) demonstrates that, for the same level of false classification of earthquakes as explosions, our methodological adjustments reduce the number of explosions missed from 25 out of 140 down to only two (Fig. 6a versus Fig. 7c).

Given our idealized intention to never mis-classify an explosion, it is worth scrutinizing these two events in more detail. The first is located close to the high-density patch of explosions on the fundamental lune in Fig. 7(c) and corresponds to the DAG-1 event. This test was a chemical explosion with very low yield (only ∼0.9 tonnes TNT equivalent) and was performed at the Nevada National Security Site on 2018 July 20 (Ichinose *et al.* 2021). In order to ensure that it is screened in as an explosion, we found that it was necessary to increase the prior probability to $\pi_1 = 0.99999999$, which increases the associated number of mis-classified earthquakes to 38 (∼3 per cent). The second mis-classified explosion is positioned on the right-hand side of Fig. 7(c) and corresponds to the 1983 May 5 Crowdie test in Nevada. As well as being of relatively low yield (∼6 kT of TNT equivalent), the associated seismic data exhibit high levels of noise and poor waveform fits during moment tensor inversion, such that the inferred solution is associated with large uncertainties (Pasyanos & Chiang 2021). While its apparent position as an extreme outlier renders it unlikely to ever be correctly classified using our current approach, a fruitful avenue for future research would be to also map each moment tensor's individual uncertainty into the discrimination probabilities. Suitable methods are currently under development to robustly constrain the uncertainties introduced into moment tensor solutions by noisy data and/or theoretical errors (e.g. imperfect knowledge of the local seismic velocity structure, Nayak & Dreger 2018; Chiang *et al.* 2018; Phạm & Tkalčić 2021; Hu *et al.* 2023) and will likely prove helpful in this endeavour.

3.3 Applicability to monitoring in other regions

In developing this moment-tensor discrimination method, we have primarily focused on the event catalogues of Pasyanos & Chiang (2021) that contain earthquakes and explosions from the western United States. It is therefore desirable to assess its success at classifying events from other locations. Whilst we expect the physical mechanisms taking place in the source region of explosions and earthquakes to remain consistent in different spatial settings, there are two aspects that may potentially affect the rate of success. First, for many regions of potential interest, ground-truth catalogues of explosions and earthquakes are not available and we are therefore limited to using population description metrics on the hypersphere obtained from the U.S. examples. Second, seismometer coverage and knowledge of the local seismic velocity structure is likely to be poorer in many regions, which feeds into larger uncertainties in the inverted moment tensors that are input into our classification scheme.

Full assessment of the impact of these issues on screening success rates will require further work, but we can gain an initial insight into the method's applicability in other regions by assessing its ability to identify known nuclear tests that have been conducted outside of the United States. We focus on published moment tensors for the six declared nuclear tests that were carried out in the DPRK between 2006 and 2017. Chiang *et al.* (2018) provide full moment tensor solutions for these six tests, plus a collapse event that occurred immediately following the 2017 test. Alvizuri & Tape (2018) provide solutions

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

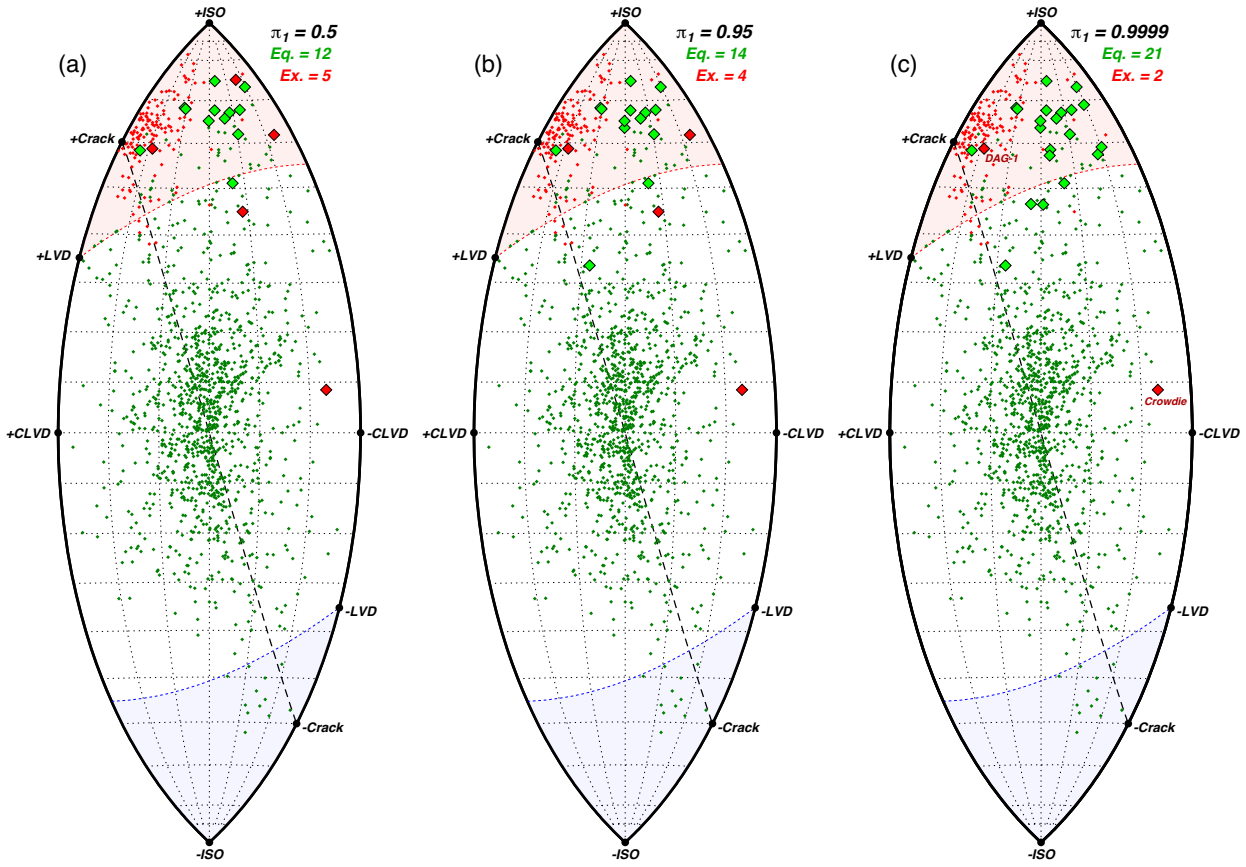

Figure 7. Mis-classification of event source type for different priors. (a) 140 explosions and 1149 earthquakes from Fig. 1(b) (Pasyanos & Chiang 2021) screened using SvMF distributions fitted to the explosions and earthquakes using a prior probability that each event was an explosion of $\pi_1 = 0.5$ (i.e. equal priors, equivalent to maximum likelihood). Large red diamonds = explosions (Ex.) that are incorrectly classified as earthquakes; large green diamonds = earthquakes (Eq.) incorrectly classified as explosions; inset lists the number of mis-classified events. (b) Same for $\pi_1 = 0.95$. (c) Same for $\pi_1 = 0.9999$, with the Crowdie and DAG-1 tests labelled. All data have been projected onto the fundamental lune of Tape & Tape (2012), which is plotted using a Hammer equal area projection.

for these same seven events plus two earthquakes that occurred in 2016 and 2017 in South Korea, while both Liu *et al.* (2018) and Hu *et al.* (2023) provide their own moment tensor solutions for the five 2009–2017 tests. Each of these studies explored different inversion methods and sources of uncertainty, resulting in variations in the optimal moment tensor solutions. Nevertheless, our method identifies all of the tests as explosions and also successfully assigns the two earthquakes to their correct category. Both versions of the collapse event end up in the earthquake category. The method therefore appears suitable for categorizing moment tensors associated with explosive events from other localities.

4 CONCLUSIONS

Seismic moment tensors provide a robust method to describe event-source characteristics and differentiate between explosions and earthquakes, rendering them an important tool for monitoring underground nuclear testing (Ford *et al.* 2020). Here, we have investigated published catalogues of known events and demonstrated that, when normalized and projected onto their six-degree unit vectors, population distributions of moment tensor solutions for explosions and earthquakes exhibit significant levels of anisotropy. We show that elliptically symmetric distribution functions, which have the property of differing levels of variance in each dimension, are a more appropriate description of event populations than the spherically symmetric (i.e. isotropic) distributions that have previously been used. We also show that describing both explosion and earthquake populations using individual probability density functions allows switching from a simple yes/no screening statistic to a maximum-likelihood approach. Comparing performance of the isotropic and elliptical probability density functions on the event catalogues of Pasyanos & Chiang (2021) yields an improvement in classification rates. For equal prior probabilities of an event belonging to the explosion or earthquake populations, our elliptical model mis-classifies only 17 events out of 1289 total, in comparison to 39 for the isotropic version. By adjusting the prior probabilities, our approach can achieve classification rates of up to 99 per cent for explosions at a cost of less than 2 per cent mis-classification of earthquakes. Quantifying the probabilities of any new event belonging to each population takes a fraction of a second and is therefore suitable for real-time deployment on moment tensor solutions.

In addition to improving classification rates for seismic moment tensors, these updates provide a number of other advantages. First, by providing a probability metric, it is possible to further identify several mis-classified events and also easily integrate the method into existing monitoring workflows that leverage multiple independent discriminants. Secondly, introduction of a prior probability factor provides a simple means to adjust the trade-off between potentially

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

missing explosions versus mis-classifying earthquakes. Finally, the scheme is flexible to the incorporation of additional event populations, providing the possibility of assessing other event types in future, such as collapses.

ACKNOWLEDGMENTS

We are grateful to Michael Pasyanos and Doug Dreger for sharing their catalogue of earthquake focal mechanisms. Editor Carl Tape, Sean Ford and two other anonymous reviewers provided helpful and encouraging feedback. The study also benefitted from interactions with Jinyin Hu, Thành-Son Phạm, Hrvoje Tkalčić, Sia Ghelichkhan and Thomas Duvernay. This research is supported by Geoscience Australia and the Australian Research Council via a Discovery Early Career Researcher Award (DE220101519) and Discovery Project (DP220102232).

DATA AVAILABILITY

Moment tensor solutions are either publicly available from cited sources or by reasonable request to Michael Pasyanos. A simple python routine `MTid` and tutorial are available at https://github.com/mjh217/MTid that allow a user to assess the most likely source type for any moment tensor using the approach described herein. The code version used to generate these results is release v2023.12 (doi:10.5281/zenodo.10408592).

REFERENCES

Alvizuri, C. & Tape, C., 2018. Full moment tensor analysis of nuclear explosions in North Korea, *Seismol. Res. Lett.,* **89,** 2139–2151.

Anderson, D.N., Fagan, D.K., Tinker, M.A., Kraft, G.D. & Hutchenson, K.D., 2007. A mathematical statistics formulation of the teleseismic explosion identification problem with multiple discriminants, *Bull. seism. Soc. Am.,* **97,** 1730–1741.

Blandford, R.R., 1977. Discrimination between earthquakes and underground explosions, *Annu. Rev. Earth planet. Sci.,* **5,** 111–122.

Bowers, D. & Selby, N.D., 2009. Forensic seismology and the Comprehensive Nuclear-Test-Ban Treaty, *Annu. Rev. Earth planet. Sci.,* **37,** 209–236.

Boyd, O.S., Dreger, D.S., Lai, V.H. & Gritto, R., 2015. A systematic analysis of seismic moment tensor at The Geysers geothermal field, California, *Bull. seism. Soc. Am.,* **105,** 2969–2986.

Brune, J.N. & Pomeroy, P.W., 1963. Surface wave radiation patterns for underground nuclear explosions and small-magnitude earthquakes, *J. geophys. Res.,* **68,** 5005–5028.

Carmichael, J.D., 2021. Hypothesis tests on Rayleigh wave radiation pattern shapes: A theoretical assessment of idealized source screening, *J. geophys. Int.,* **225,** 1653–1671.

Chiang, A., Ichinose, G.A., Dreger, D.S., Ford, S.R., Matzel, E.M., Myers, S.C. & Walter, W.R., 2018. Moment tensor source-type analysis for the Democratic People's Republic of Korea – declared nuclear explosions (2006–2017) and 3 September 2017 collapse event, *Seismol. Res. Lett.,* **89,** 2152–2165.

Dreger, D. & Woods, B., 2002. Regional distance seismic moment tensors of nuclear explosions, *Tectonophysics,* **356,** 139–156.

Dreger, D.S., Tkalčić, H. & Johnston, M., 2000. Dilational processes accompanying earthquakes in the Long Valley Caldera, *Science,* **288,** 122–125.

Evernden, J.F., Best, W.J., Pomeroy, P.W., McEvilly, T.V., Savino, J.M. & Sykes, L.R., 1971. Discrimination between small-magnitude earthquakes and explosions, *J. geophys. Res.,* **76,** 8042–8055.

Fisk, M.D., Jepsen, D. & Murphy, J.R., 2002. Experimental seismic event-screening criteria at the Prototype International Data Center, *Pure appl. Geophys.,* **159,** 865–888.

Ford, S.R., Dreger, D.S. & Walter, W.R., 2009. Identifying isotropic events using a regional moment tensor inversion, *J. geophys. Res.,* **114,** B01306. doi: 10.1029/2008JB005743.

Ford, S.R., Kraft, G.D. & Ichinose, G.A., 2020. Seismic moment tensor event screening, *J. geophys. Int.,* **221,** 77–88.

Gilbert, F., 1970. Excitation of the normal modes of the Earth by earthquake sources, *Geophys. J. R. astr. Soc.,* **22,** 223–226.

Hastie, T., Tibshirani, R. & Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* 2nd edn. Springer, New York

Hu, J., Phạm, T.-S. & Tkalčić, H., 2023. Seismic moment tensor inversion with theory errors from 2-D Earth structure: implications for the 2009–2017 DPRK nuclear blasts, *J. geophys. Int.,* **235**(3), 2035–2054.

Ichinose, G.A., Ford, S.R., Kroll, K., Dodge, D., Pyle, M., Pitarka, A. & Walter, W.R., 2021. Preliminary analysis of source physics experiment explosion-triggered microseismicity using the back-projection method, *J. geophys. Res.: Solid Earth,* **126,** e2020JB021312. doi: 10.1029/2020JB021312.

Julian, B.R., Miller, A.D. & Foulger, G.R., 1998. Non-double-couple earthquakes 1. Theory, *Rev. Geophys.,* **36,** 525–549.

Kent, J.T., 1982. The Fisher—Bingham distribution on the sphere, *J. R. Stat. Soc.: Ser. B,* **44,** 71–80.

Koper, K.D., Holt, M.M., Voyles, J.R., Burlacu, R., Pyle, M.L., Wang, R. & Schmandt, B., 2021. Discrimination of small earthquakes and buried single-fired chemical explosions at local distances (<150 km) in the western United States from comparison of local magnitude (M_L) and coda duration magnitude (M_C), *Bull. seism. Soc. Am.,* **111,** 558–570.

Liu, J., Li, L., Zahradník, J., Sokos, E. & Plicka, V., 2018. Generalized source model of the North Korea tests 2009–2017, *Seismol. Res. Lett.,* **89,** 2166–2173.

Minson, S.E. & Dreger, D.S., 2008. Stable inversions for complete moment tensors, *J. geophys. Int.,* **174,** 585–592.

Nayak, A. & Dreger, D.S., 2018. Source inversion of seismic events associated with the sinkhole at Napoleonville salt dome, Louisiana using a 3-D velocity model, *J. geophys. Int.,* **214,** 1808–1829.

Paine, P.J., Preston, S.P., Tsagris, M. & Wood, A. T.A., 2018. An elliptically symmetric angular Gaussian distribution, *Stat. Comput.,* **28,** 689–697.

Pasyanos, M.E. & Chiang, A., 2021. Full moment tensor solutions of U.S. underground nuclear tests for event screening and yield estimation, *Bull. seism. Soc. Am.,* **112,** 538–552.

Phạm, T.-S. & Tkalčić, H., 2021. Toward improving point-source moment-tensor inference by incorporating 1D Earth model's uncertainty: Implications for the Long Valley Caldera earthquakes, *J. geophys. Res.: Solid Earth,* **126,** e2021JB022477. doi: 10.1029/2021JB022477.

Pomeroy, P.W., Best, W.J. & McEvilly, T.V., 1982. Test Ban Treaty verification with regional data – a review, *Bull. seism. Soc. Am.,* **72,** S89–S129.

Rivest, L.-P., 1984. On the information matrix for symmetric distributions on the hypersphere, *Ann. Stat.,* **12,** 1085–1089.

Scealy, J.L. & Wood, A. T.A., 2019. Scaled von Mises–Fisher distributions and regression models for paleomagnetic directional data, *J. Am. Stat. Assoc.,* **114,** 1547–1560.

Scealy, J.L., Heslop, D., Liu, J. & Wood, A. T.A., 2022. Directions old and new: palaeomagnetism and Fisher (1953) meet modern statistics, *Int. Stat. Rev.,* **90,** 237–258.

Selby, N.D., Marshall, P.D. & Bowers, D., 2012. m_b:M_s event screening revisited, *Bull. seism. Soc. Am.,* **102,** 88–97.

Silver, P.G. & Jordan, T.H., 1982. Optimal estimation of scalar seismic moment, *Geophys. J. R. astr. Soc.,* **70,** 755–787.

Stevens, J.L. & O'Brien, M., 2018. 3D nonlinear calculation of the 2017 North Korean nuclear test, *Seismol. Res. Lett.,* **89,** 2068–2077.

Tape, W. & Tape, C., 2012. A geometric setting for moment tensors, *J. geophys. Int.,* **190,** 476–498.

Tape, W. & Tape, C., 2015. A uniform parametrization of moment tensors, *J. geophys. Int.,* **202,** 2074–2081.

Wallace, T.C., 1991. Body wave observations of tectonic release, in *Explosion Source Phenomenology,* chap. 13, pp. 161–170, eds Taylor, S.R., Patton, H.J. & Richards, P.G., American Geophysical Union, Washington, D.C.

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

Walter, W.R., Mayeda, K.M. & Patton, H.J., 1995. Phase and spectral ratio discrimination between NTS earthquakes and explosions. Part I: Empirical observations, *Bull. seism. Soc. Am.,* **85,** 1050–1067.

Wood, A.T.A., 1994. Simulation of the von Mises-Fisher distribution, *Commun. Stat.—Simul. Comput.,* **23,** 157–164.

Downloaded from https://academic.oup.com/gji/article/237/1/1/7597260 by guest on 07 February 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of The Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.