

Quantifying the similarity of seismic polarizations

Joshua P. Jones,¹ David W. Eaton² and Enrico Caffagni³¹220 SE 151st Ave, Portland, OR 97233, USA. E-mail: highly.creative.pseudonym@gmail.com²Department of Geoscience, University of Calgary, Calgary, AB T2N 1N4, Canada³Department of Meteorology and Geophysics, University of Vienna, Althanstraße 14 (UZA II), 1090 Vienna, Austria

Accepted 2015 November 9. Received 2015 October 22; in original form 2014 December 2

SUMMARY

Assessing the similarities of seismic attributes can help identify tremor, low signal-to-noise (S/N) signals and converted or reflected phases, in addition to diagnosing site noise and sensor misalignment in arrays. Polarization analysis is a widely accepted method for studying the orientation and directional characteristics of seismic phases via computed attributes, but similarity is ordinarily discussed using qualitative comparisons with reference values or known seismic sources. Here we introduce a technique for quantitative polarization similarity that uses weighted histograms computed in short, overlapping time windows, drawing on methods adapted from the image processing and computer vision literature. Our method accounts for ambiguity in azimuth and incidence angle and variations in S/N ratio. Measuring polarization similarity allows easy identification of site noise and sensor misalignment and can help identify coherent noise and emergent or low S/N phase arrivals. Dissimilar azimuths during phase arrivals indicate misaligned horizontal components, dissimilar incidence angles during phase arrivals indicate misaligned vertical components and dissimilar linear polarization may indicate a secondary noise source. Using records of the $M_w = 8.3$ Sea of Okhotsk earthquake, from Canadian National Seismic Network broad-band sensors in British Columbia and Yukon Territory, Canada, and a vertical borehole array at Hoadley gas field, central Alberta, Canada, we demonstrate that our method is robust to station spacing. Discrete wavelet analysis extends polarization similarity to the time–frequency domain in a straightforward way. Time–frequency polarization similarities of borehole data suggest that a coherent noise source may have persisted above 8 Hz several months after peak resource extraction from a ‘flowback’ type hydraulic fracture.

Key words: Time-series analysis; Wavelet transform; Persistence, memory, correlations, clustering; Body waves.

1 INTRODUCTION

Polarization analysis has been used to characterize seismic data for decades. Some of the earliest quantitative studies of particle motion were performed by Montalbetti & Kanasevich (1970), Vidale (1986) and René *et al.* (1986). Decades later, however, no technique exists that effectively quantifies the notion of ‘similar polarizations’. Existing techniques for polarization analysis often compare polarizations with predicted values from located earthquakes to establish empirical misfit distributions (e.g. Jurkevics 1988; De Meersman *et al.* 2006). In monitoring experiments where multiple seismic sources are simultaneously active, researchers compare polarization with expected values from different plausible physical sources (e.g. Acernese *et al.* 2004; Jones *et al.* 2012b; Das & Zoback 2013b; Tary *et al.* 2014), yet these comparisons are qualitative and sometimes only graphical.

In recent years, blind source separation (BSS) techniques have been used to extract similarly polarized signals from narrow aperture seismic arrays; examples in the literature use Principal Components Analysis (Jones *et al.* 2012a,b), Singular Value Decomposition (De Meersman *et al.* 2006), and Independent Component Analysis (Acernese *et al.* 2004). Yet BSS can only assess similarity using proxy measures to compare each sensor with a common (extracted) spatio-temporal pattern, such as eigenvector loadings at each sensor (Jones *et al.* 2012a), or deviations from a calculated reference value (De Meersman *et al.* 2006). The latter is also a traditional means of assessing accuracy when studying computed values of polarization attributes (e.g. Vidale 1986; Park *et al.* 1987; Jurkevics 1988; Pinnegar 2006).

In this work, we introduce a quantitative measure for polarization similarity based on histogram distance metrics adapted from the image processing literature. The proposed method accounts for

Table 1. Polarization attributes used in this work.

Name	Symbol and formula	Range	Reference
Azimuth	$\theta = \tan^{-1} \left(\frac{\text{Re}(v_{13})}{\text{Re}(v_{12})} \right)$	$(-90^\circ, 90^\circ]$	Vidale (1986)
Ellipticity	$\eta = \frac{\sqrt{(1-\text{Re}(v_1^2))}}{\text{Re}(v_1)}$	$[0, 1]$	Vidale (1986)
Incidence	$\phi = \tan^{-1} \left(\frac{\sqrt{\text{Re}(v_{12}^2) + \text{Re}(v_{13}^2)}}{\text{Re}(v_{11})} \right)$	$(-90^\circ, 90^\circ]$	Vidale (1986) ^a
Planarity	$\nu = 1 - \frac{2\lambda_3}{\lambda_1 + \lambda_2}$	$[0, 1]$	Jurkevics (1988)
Rectilinearity	$\rho = 1 - \frac{\lambda_2 + \lambda_3}{\lambda_1}$	$[0, 1]$	Jurkevics (1988)

^aVidale (1986) uses dip, $\delta = \tan^{-1} ((\text{Re}(v_{12}^2) + \text{Re}(v_{13}^2))^{0.5} / \text{Re}(v_{11}))$, which has the opposite sense of ϕ .

some degree of scatter by computing the cost of transforming one histogram into another, and is easily modified for periodic (angular) polarization attributes. Weighting histograms by seismic energy can suppress noise and allow characterization of emergent arrivals. Renormalizing histograms prior to the similarity calculation eliminates the need to determine scalar site amplifications.

2 THEORY: SEISMIC POLARIZATION

We start with a brief review of quantitative polarization analysis. For a single-channel digital seismogram \mathbf{x} , the analytic extension of the signal is

$$\tilde{\mathbf{x}} = \mathbf{x} + ih(\mathbf{x}), \quad (1)$$

where h denotes the Hilbert transform (Kanasewich 1981). For three-component data from multiple sensors, let k denote sensor index and assume a right-handed Z, N, E coordinate system (with positive Z indicating downward motion). Let $\mathbf{X}_k = [\mathbf{x}_{zk} \ \mathbf{x}_{nk} \ \mathbf{x}_{ek}]$ denote a three-component seismogram. Polarization can be estimated by diagonalizing the complex covariance matrix

$$\mathbf{C} = \begin{bmatrix} \tilde{\mathbf{x}}_z \cdot \tilde{\mathbf{x}}_z^* & \tilde{\mathbf{x}}_z \cdot \tilde{\mathbf{x}}_n^* & \tilde{\mathbf{x}}_z \cdot \tilde{\mathbf{x}}_e^* \\ \tilde{\mathbf{x}}_n \cdot \tilde{\mathbf{x}}_z^* & \tilde{\mathbf{x}}_n \cdot \tilde{\mathbf{x}}_n^* & \tilde{\mathbf{x}}_n \cdot \tilde{\mathbf{x}}_e^* \\ \tilde{\mathbf{x}}_e \cdot \tilde{\mathbf{x}}_z^* & \tilde{\mathbf{x}}_e \cdot \tilde{\mathbf{x}}_n^* & \tilde{\mathbf{x}}_e \cdot \tilde{\mathbf{x}}_e^* \end{bmatrix} \quad (2)$$

where $*$ denotes the complex conjugate (Kanasewich 1981; Vidale 1986; D'Auria *et al.* 2010); the complex time–frequency transform in Kulesh *et al.* (2007) yields an equivalent formulation. The system

$$\begin{bmatrix} \tilde{\mathbf{x}}_z \\ \tilde{\mathbf{x}}_n \\ \tilde{\mathbf{x}}_e \end{bmatrix} \mathbf{C} - \Lambda \mathbf{I} = 0 \quad (3)$$

is solved to obtain real eigenvalues λ_i ($i = 1, 2, 3$) and complex eigenvectors \mathbf{v}_i . Each λ_i gives the relative strength of the polarization vector and the corresponding \mathbf{v}_i give their orientation with respect to the data coordinate system (Jurkevics 1988). In order to investigate elliptical polarizations, Vidale (1986) rotated the principal eigenvector \mathbf{v}_1 of \mathbf{C} into the complex phase angle ψ that maximizes the length of $\text{Re}(\mathbf{v}_1)$.

A number of attributes can be computed from λ , ψ and \mathbf{v}_1 . As a shorthand convention we will use the generic variable α to refer to a nonspecific polarization attribute. The attributes used in this work are given in Table 1, though any bounded attribute computed by equivalent techniques (e.g. René *et al.* 1986; Kulesh *et al.* 2007) could be analysed identically.

The measures used in this work have straightforward physical interpretations. η measures elliptical polarization, with 0 indicating a

linearly polarized wavefield and 1 indicating spherical polarization. θ and ϕ are horizontal and vertical angular measures, respectively. Rectilinearity (ρ) and planarity (ν) measure the linear and planar polarization of the wavefield, respectively, where a high ρ represents a linearly polarized wavefield.

2.1 Instantaneous versus averaged polarization

As described in D'Auria *et al.* (2010) and Vidale (1986), evaluating \mathbf{C} at a single time t yields one non-null eigenvalue λ_1 with corresponding (normalized) principal eigenvector

$$\mathbf{v}_{1t} = \begin{bmatrix} \frac{\tilde{x}_{zt}}{\tilde{x}_{et}} & \frac{\tilde{x}_{nt}}{\tilde{x}_{et}} & 1 \end{bmatrix}. \quad (4)$$

Of the attributes α listed in Table 1, ellipticity (η), azimuth (θ) and incidence angle (ϕ) are sometimes called ‘instantaneous’ because they can be computed directly from eq. (4) after finding ψ (Vidale 1986). ν and ρ are termed ‘averaged’ quantities in that \mathbf{C} in eq. (2) must be averaged over a window at least 3 points long to yield three non-null eigenvalues λ_i . In practice, averaging is often used even for instantaneous measures to stabilize the calculation.

3 METHOD

A measure of similarity for seismic attributes, and particularly for seismic polarization, would be a useful tool for addressing a number of problems: potential examples include sensor alignment in arrays, empirical phase separation, and determining the repetition intervals (and hence constraining source durations) for quasi-continuous signals formed by superposed, discrete sources. This task is difficult for seismic polarization for two reasons: not only are angular attributes periodic, but in a heterogeneous environment, small changes in path geometry can yield significant changes in attribute values. Different seismic phases can show systematic changes in every attribute across an array while individual measures vary considerably (e.g. Jurkevics 1988, figs 6–8). This may be due to sensor misalignment but is more commonly a consequence of seismic path effects.

3.1 Histogram distances

Time series of polarization attributes are extremely difficult to compare due to both the 180° ambiguity in θ and ϕ and the effects of scattering on all measures. However, histogram distances are commonly used to compare feature populations in other fields, notably image processing and computer vision; a few prominent methods from the recent literature are Pele & Werman (2010), Ling & Okada

(2007), Rubner *et al.* (2000) and Hafner *et al.* (1995). In this work, we measure polarization similarity via the proxy measure of histogram distances computed from time windows of three-component seismograms.

Let H denote a histogram; let $H_{ik\alpha}$ denote the i th bin of histogram H at station k for attribute α . For instantaneous attributes like those of Vidale (1986), a natural approach is to form weighted histograms, where each bin i is filled with the summed energy of all data whose α lie within the bin's edges. Hence, each bin is given by

$$H_{aki}^{(V)} = \sum_{t:\alpha_k(t)\in i} x_{zk}^2(t) + x_{nk}^2(t) + x_{ek}^2(t). \quad (5)$$

When an attribute α is instead averaged in a short time window, as with the planarity and rectilinearity of Jurkevics (1988), the equivalent weighting scheme is to sum $\text{Tr}(\mathbf{C})$ from eq. (2),

$$H_{aki}^{(J)} = \sum_{t:\alpha_k(t)\in i} \text{Tr}(\mathbf{C}_{kt}). \quad (6)$$

We will use these weighting schemes to construct histograms of polarization attributes in the rest of this work.

3.2 Choice of distance metric

It remains to choose an appropriate distance metric for polarization histograms. In problems where features (attributes) being studied are neither periodic nor circularly shifted within their domain, similarity can be described with simple measures that compare each bin in H_1 with the same bin in H_2 ; the simplified notation arises because we are considering a single hypothetical attribute. The L_1 and L_2 norms are both examples of such 'bin-to-bin' distance metrics. Another well-known example is the ' χ^2 distance' of Snedecor & Cochran (1967),

$$D^{(\chi^2)} = \frac{1}{2} \sum_{i=1}^{N_i} \frac{(H_{1i} - H_{2i})^2}{H_{1i} + H_{2i} + \epsilon}, \quad (7)$$

where ϵ is a small corrective term that stabilizes the calculation for sparse histograms. Although based on the Pearson's χ^2 test statistic, the denominator of eq. (7) differs to satisfy the symmetry property $D(H_1, H_2) = D(H_2, H_1)$ (Munkres 1975).

Although well-suited to problems in other fields, distance metrics like $D^{(\chi^2)}$ are poorly suited to characterize seismic attributes because each attribute in Table 1 is subject to slight absolute shifts from noise and local heterogeneities. It follows from basic considerations that variations in angular attributes increase as station spacing increases. Thus, a robust way to compensate for this is to adopt a distance metric that takes the resultant bin shifts into account.

Some histogram distance metrics calculate the minimum cost of transforming one distribution into another. The canonical example of this class of metric is the Earth mover's distance (EMD) or Wasserstein metric (Rubner *et al.* 2000), adapted from the fundamental transportation theory work of Monge (1781). However, in a recent paper by Pele & Werman (2010), a new metric was introduced that outperforms EMD and its variants in both speed and accuracy: the Quadratic Chi (QC) distance,

$$D^{(QC)} = \left(\sum_{ij} \left(\frac{H_{1i} - H_{2i}}{(\sum_c (H_{1c} + H_{2c}) A_{ci})^m} \right) \left(\frac{H_{1j} - H_{2j}}{(\sum_c (H_{1c} + H_{2c}) A_{cj})^m} \right) A_{ij} \right)^{\frac{1}{2}}, \quad (8)$$

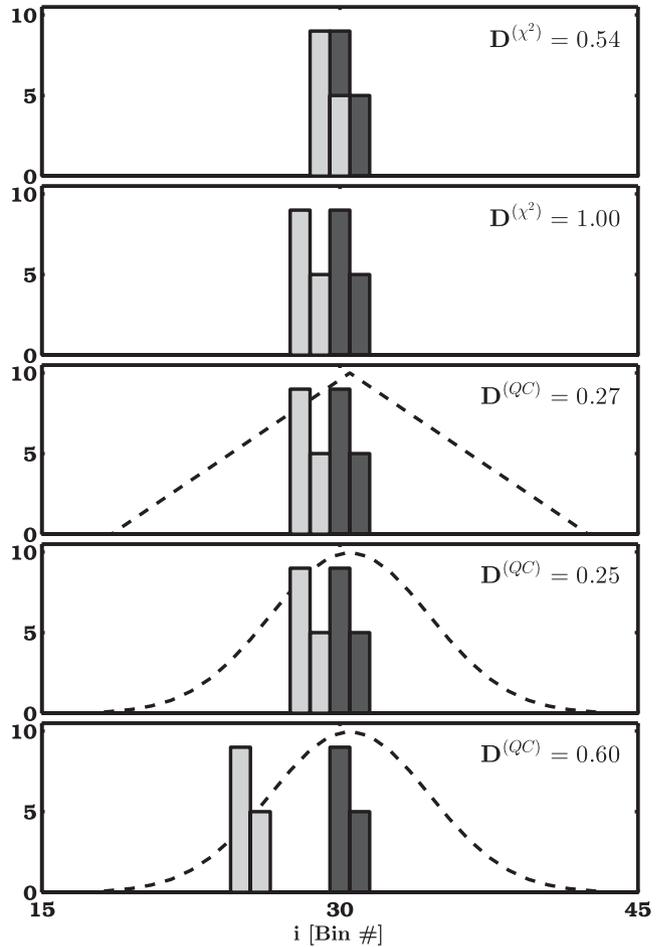


Figure 1. Distances between two fictitious populations. Reference and test populations (dark and light grey, respectively) are divided into $N_i = 60$ bin histograms. Distance values appear at the upper right of each plot. The superscript (χ^2) indicates D is calculated using the distance metric of Snedecor & Cochran (1967), while (QC) indicates that of Pele & Werman (2010). In the lower three plots, the bin distance matrix \mathbf{A} uses $\tau = 12$; its form is indicated by a dashed line centred on $i = 30$, i.e. row A_{30j} of \mathbf{A} .

where m is a normalization factor ($0 \leq m < 1$) and \mathbf{A} is a bin separation matrix (described below). In this work we strictly set $m = 0.5$; our choice of \mathbf{A} will be discussed in the next section.

An appealing property of eq. (8) is that normalization is straightforward: with $m = 0.5$ and histograms renormalized s.t. $\sum_i H_{1i} = \sum_i H_{2i} = 0.5$, $\max(D^{(QC)}) \sim 1$. As histograms become more uniformly distributed, $\max(D^{(QC)})$ decreases. However, as we demonstrate below, when polarization attributes are computed in a time window around an impulsive phase arrival, their histograms are generally sparse and well-peaked.

3.2.1 The bin separation matrix

The bin separation matrix \mathbf{A} in eq. (8) increases the distance between two histograms using a cost functional that penalizes moving material between bins. The choice of cost functional is somewhat arbitrary; Pele & Werman (2010) recommend measures adapted to the data being studied. In this section, we introduce and motivate a bin separation matrix suitable for studying seismic polarization.

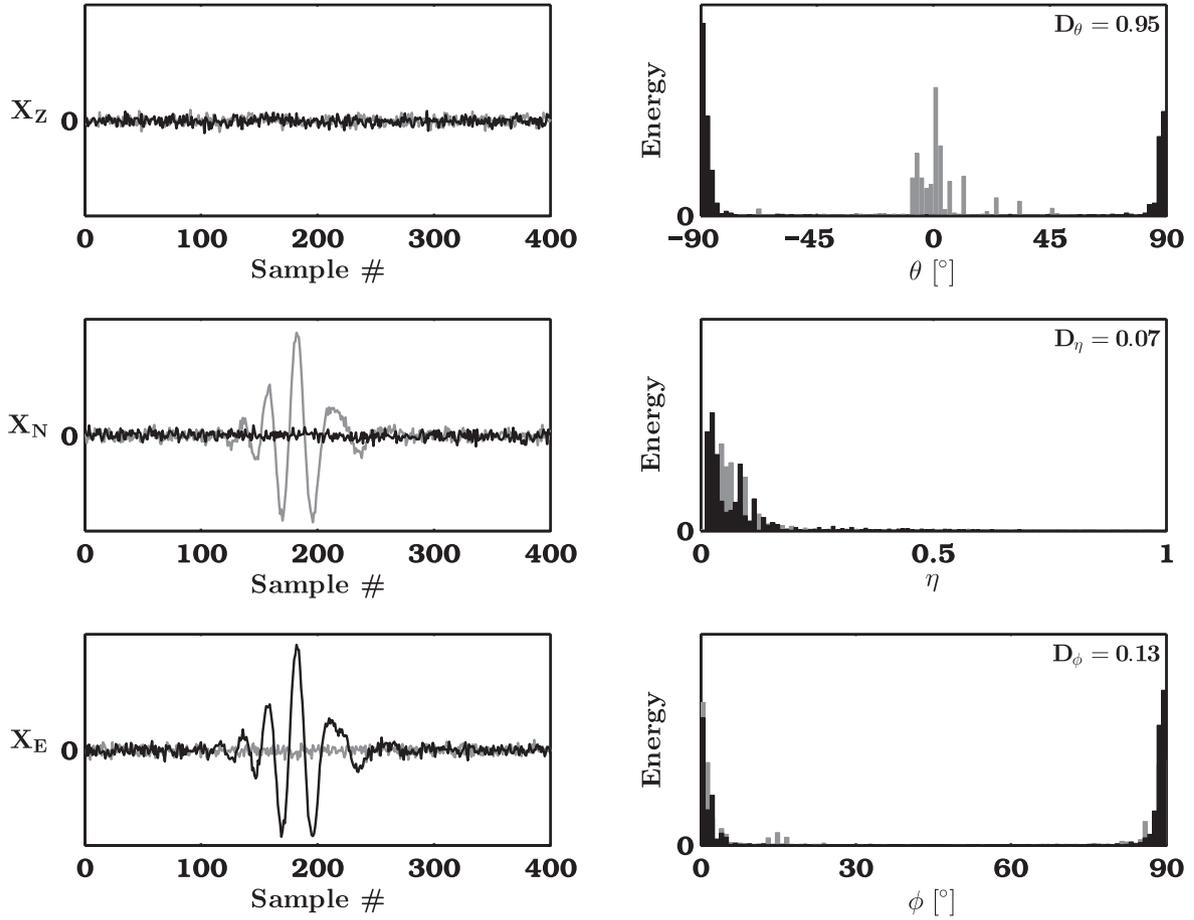


Figure 2. Sample synthetic data and corresponding polarization histograms for azimuth (θ , top right), ellipticity (η , middle right), and incidence (ϕ , bottom right). Grey lines and bars are generated from a noise-free reference signal with polarization $\theta = 0^\circ$, $\eta = 0$, $\phi = 90^\circ$. Black lines and bars correspond to a test signal with polarization $\theta = 90^\circ$, $\eta = 0$, $\phi = 90^\circ$. Additive Gaussian noise is generated independently for each channel at S/N = 10 dB. Population distances computed via eq. (8) appear at the upper right of each histogram.

The simplest form of \mathbf{A} is the identity matrix. For renormalized histograms with N_i bins, letting $\mathbf{A} = \mathbf{I}$ reduces $D^{(\text{QC})}$ to $\sqrt{2D(\chi^2)}$. Another simple form of \mathbf{A} scales A_{ij} linearly with bin separation $|i - j|$, i.e.

$$A_{ij} = \max\left(0, 1 - \frac{\tau|i - j|}{N_i}\right), \quad (9)$$

where N_i is number of bins in the histogram. The threshold τ is introduced so that $A_{ij} = 0$ for $|i - j| \geq \tau$.

For histograms of seismic attributes computed around a phase arrival, one might reasonably expect a roughly Gaussian distribution, rather than a linear fall-off. Thus, a more suitable expression here is a Gaussian, where we choose τ based on bin spacing and its relation to the spread in polarization attributes. This can be written explicitly, with $\mu = i$ and $\sigma = \tau/3$, as

$$A_{ij} = \begin{cases} \sqrt{\frac{3}{2\pi\tau}} \exp\left(-\frac{9(j-i)^2}{2\tau}\right) & \max(1, \lfloor i - \tau + 1 \rfloor) \\ & \leq i \leq \min(N_i, \lceil i + \tau - 1 \rceil) \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

This expression is easily periodized for θ, ϕ by relaxing the first case restriction to $\min(|i - j|, N_i - |i - j|) \leq 3\sigma$.

3.2.2 A simple example

Let us illustrate how these distance metrics behave with a ‘toy’ example, consisting of two simple histograms. The plots in Fig. 1 show two fictitious populations, P_1 (dark grey) and P_2 (light grey) divided into $N_i = 60$ bins. In all plots, the second population is shifted relative to the first.

In the top plot of Fig. 1, P_2 is shifted from P_1 by a single bin. If these histograms were computed from polarization attributes, the two populations would have an absolute shift of 3.0° (in θ or ϕ) or $1.7e-2$ (in η, ρ , or ν). The χ^2 distance between populations ($D = 0.54$) suggests little similarity. Because the two populations have only two non-empty bins, a shift of two bins or more increases the χ^2 distance to 1.0 (second plot of Fig. 1). Thus, in this example, χ^2 distance is maximized for a polarization shift of 6.0° in θ or ϕ (or $3.0e-2$ in η, ρ , or ν). A stringent distance metric that implicitly depends on number of non-empty bins is of limited use.

We now recalculate distance with eq. (8). In the lower plots of Fig. 1, we compute $D^{(\text{QC})}$ from eq. (8), setting $\tau = 12$. \mathbf{A} is computed using eqs (9) and (10) in the third and fourth plots, respectively. Using eq. (8) and the periodized version of eq. (10), two populations of polarizations become dissimilar ($D > 0.50$) when $\min(|i - j|, N_i - |i - j|) \geq 5$. (bottom plot). This bin shift corresponds to two angular polarizations shifted 15° with respect to one another.

In the rest of this work, we will use the Gaussian expression in eq. (10) to compute the bin separation matrix and eqs (9) and (8) to compute histogram distances. We will omit the superscript ^(QC) from D_α hereafter.

3.3 Attribute similarity

A static measure of polarization similarity can be defined directly from eq. (8) as

$$S_\alpha^{(s)} = 1 - 2D_\alpha, \quad (11)$$

where the factor of 2 gives S_α the approximate range and sense of a cross-correlation coefficient. However, it may be more useful to define an adaptive measure that takes the expected range of D_α into account,

$$S_\alpha^{(a)} = \min\left(1, 1 - 2 \min\left(1, \frac{D_\alpha - D_L}{D_U - D_L}\right)\right), \quad (12)$$

where D_U and D_L are empirical bounds on D_α that depend on S/N ratio. In the next section, we will establish a range for D_U and D_L using Monte-Carlo simulations of a synthetic signal with additive Gaussian noise.

3.4 Testing

Before we introduce a similarity measure, we must investigate how histogram distance behaves in a controlled way. At issue is the expected range of D_α in eq. (8) and whether averaging polarization attributes is necessary.

In this section, we explore these topics using Monte-Carlo simulations of three-component data. We generate our data as a noise-free signal with additive Gaussian noise, i.e.

$$\mathbf{X} = \mathbf{S} + \mathbf{N}, \quad (13)$$

where \mathbf{S} is a uniformly polarized seismic phase. The signal is an upsampled, zero-padded length 20 Daubechies wavelet filter (Daubechies 1992), rotated into the desired θ, ϕ . To vary η in a controlled way, we replace $\tilde{\mathbf{S}}$ with $\text{Re}(\tilde{\mathbf{S}}) \cos \psi + \text{Im}(\tilde{\mathbf{S}}) \sin \psi$, where the tilde denotes the use of eq. (1) and $\psi = (\pi/2)\eta$; this last relationship follows from the definition in Vidale (1986). For these tests, the subscript 1 will denote reference data with a fixed polarization and 2 will denote test data rotated with respect to the reference data.

3.4.1 Histogram distance range

While our choice of distance metric is bounded by [0, 1], renormalization means that two histograms of a uniformly distributed attribute α are identical ($D_\alpha = 0$). Thus, it is necessary to investigate how signal-to-noise ratio (S/N) affects the range of D_α for each parameter in Table 1.

For these tests, S/N is measured in dB, i.e. $S/N = 20 \log_{10} \|\mathbf{S}\|/\|\mathbf{N}\|$. We perform Monte-Carlo simulations by repeatedly generating pairs of synthetic data sets, computing histograms for polarization attributes, and measuring the distance between the reference and test populations. The random variable for each iteration is the additive Gaussian noise. In the first set of tests, $\mathbf{S}_1, \mathbf{S}_2$ are identically polarized ($|\Delta\alpha| = 0$ or 0° , as appropriate). Measuring D_α for two identically polarized signals with different additive noise establishes an empirical lower bound of eq. (8) as a function of S/N. In the second set of tests, the polarizations of $\mathbf{S}_1, \mathbf{S}_2$ are maximally dissimilar ($|\Delta\alpha| = 0$ or 90°), enabling us to establish an empirical upper bound for eq. (8). In both sets of tests, S/N varies from

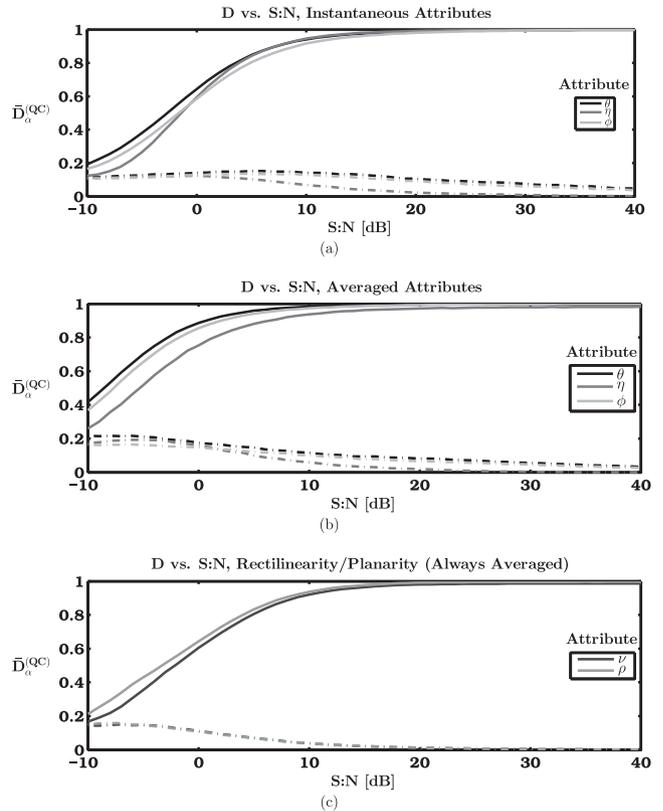


Figure 3. Minimum and maximum distances between polarization attribute histograms for synthetic data, as a function of S/N ratio in dB. Each α , S/N point averages 1000 simulations. Broken lines correspond to average distances of identically polarized signals ($\Delta\alpha = 0$). Solid lines represent average distances of maximally dissimilar polarizations ($|\Delta\alpha| = 1$ or 90° , as appropriate). (a) Distances for θ, η and ϕ when attributes are computed instantaneously using eq. (4). (b) Distances for θ, η and ϕ computed with 11-point averages. (c) Distances for ν, ρ using 11-point averages.

–10 to +40 dB in 1 dB increments, controlled using the envelopes ($E(\mathbf{x}) = |\tilde{\mathbf{x}}|$) of \mathbf{S} and \mathbf{N} . 1000 trials are conducted for each α at each S/N ratio. Histogram distances are computed using eqs (8) and (10) with $N_i = 100$ and $\tau = N_i/6$.

Fig. 2 shows a sample trial at S/N = 10 dB. Reference data \mathbf{X}_1 are grey, test data \mathbf{X}_2 are black. \mathbf{S}_1 has azimuth $\theta = 0^\circ$, \mathbf{S}_2 has $\theta = 90^\circ$. Both synthetic phase arrivals have incidence $\phi = 90^\circ$, ellipticity $\eta = 0$. As expected, only the distance between θ histograms (populations) is large.

Fig. 3 plots the results of these tests. In Fig. 3(a), distances are determined from histograms of instantaneous attributes; that is, $\theta_{kt}, \eta_{kt}, \& \phi_{kt}$ are each computed at a single time t to form each histogram $H_{\alpha k}$. In Fig. 3(b), these attributes are instead computed using an 11-point moving average. Fig. 3(c) shows the corresponding range of D_ν, D_ρ using 11-point moving averages. Comparing Figs 3(a) and (b), for $S/N \leq \sim 12$ dB, averaging increases the minimum and maximum of D_α , but the net effect is increased separation of the upper and lower bounds. At higher S/N ratios, averaging reduces the expected minimum distance between identical populations.

3.4.2 Static versus adaptive similarity

To determine which similarity measure best characterizes seismic polarization, we generate empirical fall-off curves for eqs (11) and (12) using additional Monte-Carlo simulations. Testing details are

Table 2. Testing parameters for Monte-Carlo simulations of polarization similarity using synthetic three-component data. Second column gives approximate peak of each attribute α with the reference polarization. Third column gives histogram bin width. Fourth and fifth columns give testing range and increment, respectively. Sixth column gives fixed attributes α for each set of tests.

Attribute	Ref.	Bin	Test range	Incr.	Fixed attributes
Azimuth (θ)	$\theta = 0^\circ$	1.8°	$0-30^\circ$	1°	$\phi = 90^\circ, \eta = 0$
Ellipticity (η)	$\eta = 0$	0.01	$0.00-0.30$	0.01	$\theta = 45^\circ, \phi = 45^\circ$
Incidence (ϕ)	$\phi = 90^\circ$	1.8°	$0-30^\circ$	0.5°	$\theta = 0^\circ, \eta = 0$

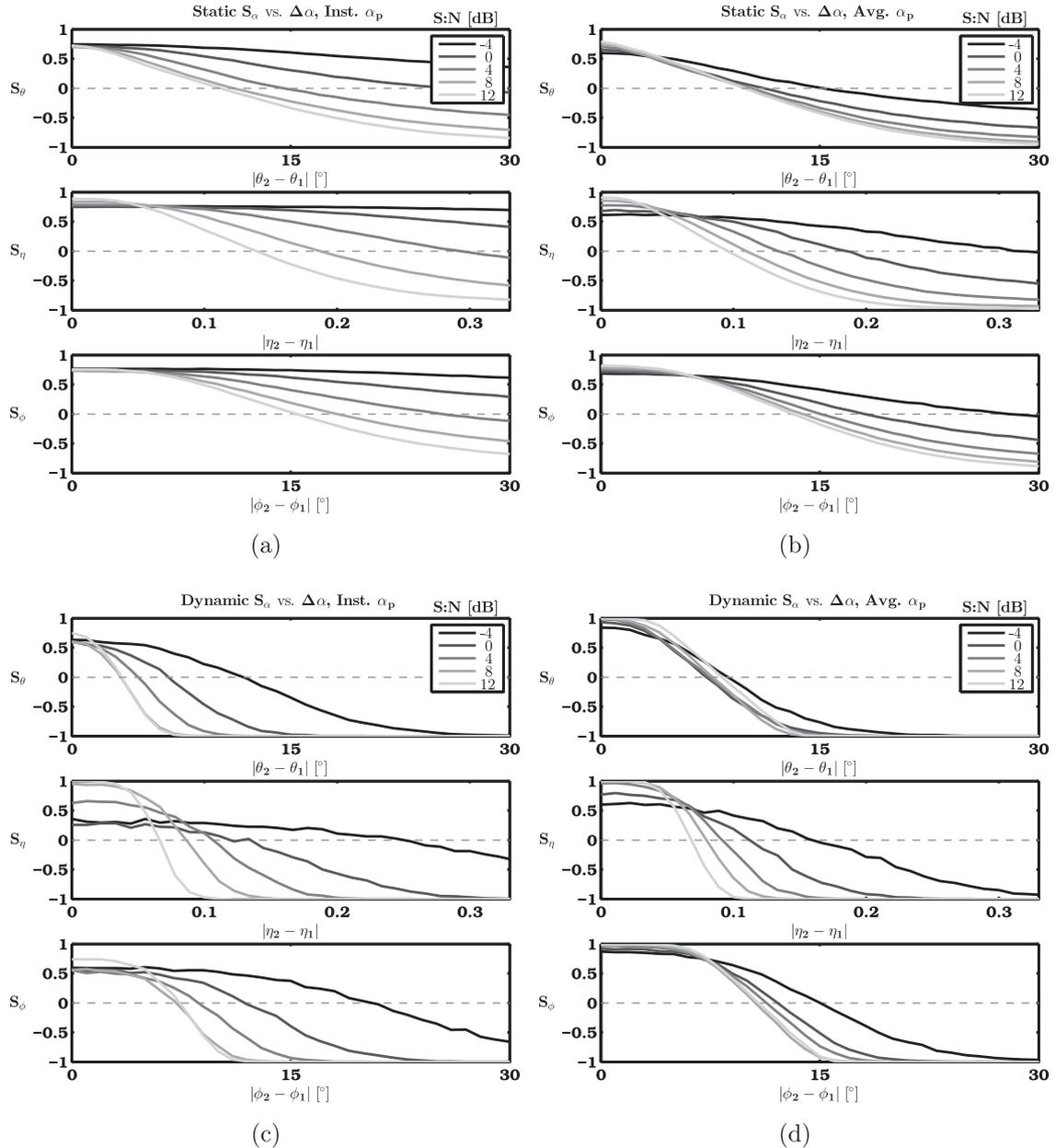


Figure 4. Similarity fall-off curves, as a function of separation between polarization peaks for two distributions computed from synthetic data. Contours correspond to different S/N ratios in dB. (a) Curves from instantaneous attributes evaluated with the static similarity measure in eq. (11). (b) Curves from attributes computed with 11-point averages using static similarity. (c) Curves from instantaneous attributes using the adaptive similarity measure in eq. (12). (d) Curves from attributes computed with 11-point averages using adaptive similarity.

Table 3. Data preprocessing guide for figures. Motivation for preprocessing differences is given in the text. Third column lists new instrument critical frequency f_c , fourth column gives new sampling frequency f_s , and fifth column gives the corner frequencies of a 2-pole zero-phase Butterworth bandpass filter; all are in Hz.

Earthquake	Array	f_c	f_s	Passband	Figs
$M_N = 3.1$, SE of Banff, AB	Hoadley, AB	3.0	100	1.0–20.0	5, 6, 7, 9
$M_c = 8.3$, Sea of Okhotsk	Hoadley, AB	0.2	50	0.04–4.5	10, 11
$M_c = 8.3$, Sea of Okhotsk	CNSN, BC & YT	$\frac{1}{120}$	50	0.04–4.5	12
$M_N = 3.1$, SE of Banff, AB	Hoadley, AB	3.0	32	Unfiltered	13
$M_c = 8.3$, Sea of Okhotsk	Hoadley, AB	0.2	32	Unfiltered	14

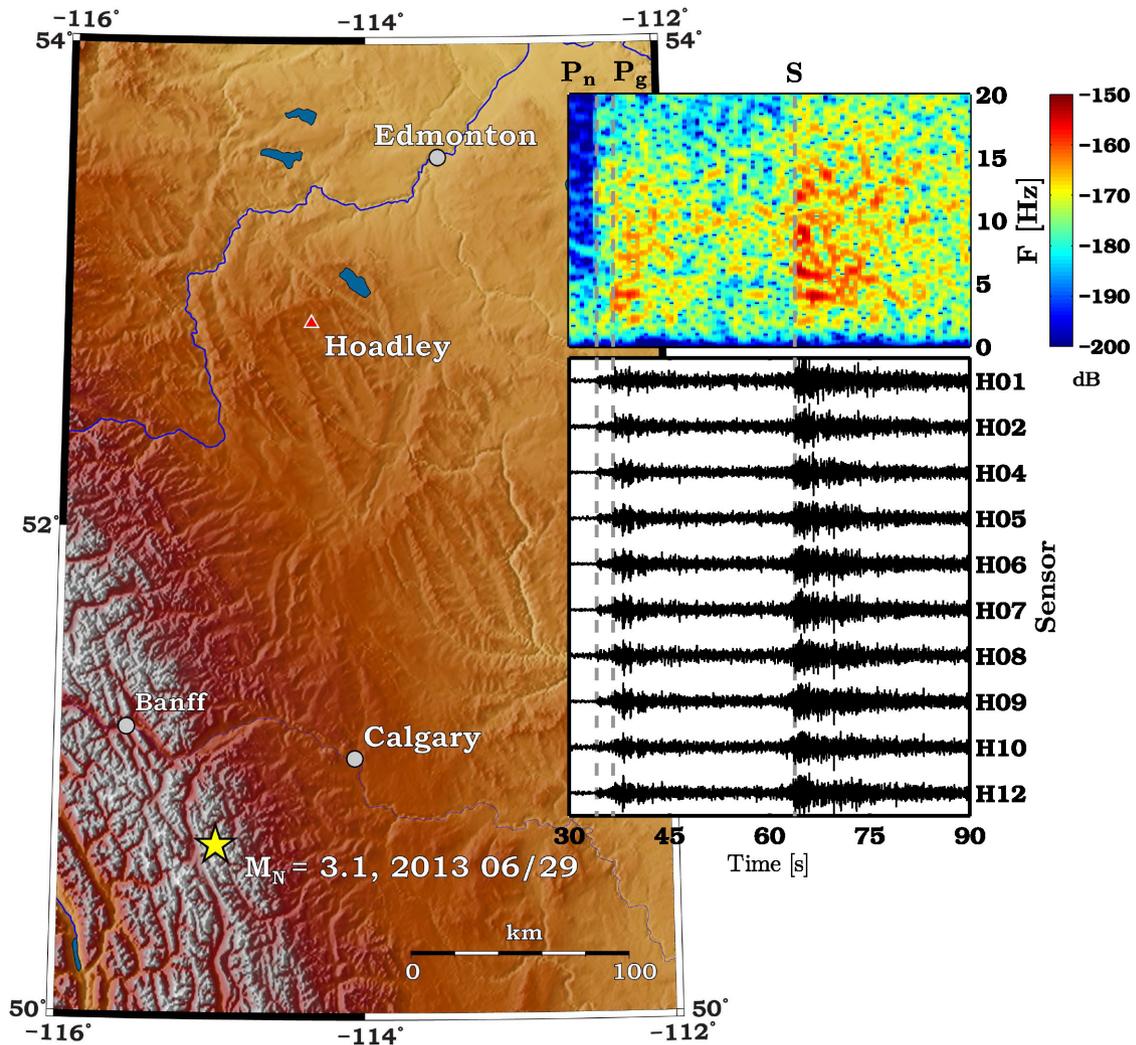


Figure 5. A regional earthquake near Banff, AB (Canada), recorded by a 12-station vertical borehole monitoring array near Hoadley, AB (red triangle). Yellow star indicates epicentre. Upper inset shows spectrogram for sensor 1 vertical component data with scaling in dB. Lower inset shows normalized, time-aligned vertical component velocity seismograms for all operational geophones, sorted by sensor. P_n , P_g and S phase arrivals are indicated with dashed lines and labels. Trace data are labelled with corresponding station numbers. Times are relative to earthquake origin.

identical to the previous section, except for the following: here, the reference data X_1 use a fixed polarization, while test data X_2 are rotated over a range of values for each parameter (Table 2). Each set of trials varies one attribute α in the test data while the other two are held fixed. 1000 trials are conducted for each α at each incremental rotation as S/N varies from -4 to 12 dB in 4 dB increments.

Fig. 4 plots the resultant curves. These can be easily interpreted; for example, at $S/N = 12$ dB, two seismic phases rotated $\theta = 30^\circ$ are dissimilar according to both similarity measures. However, comparing Figs 4(a) and (c) with Figs 4(b) and (d), the combination of averaging and adaptive similarity produces extremely uniform fall-off curves. In fact, the values in Fig. 4(d) suggest that using eq. (12) with averaged values produces such uniform results that similar

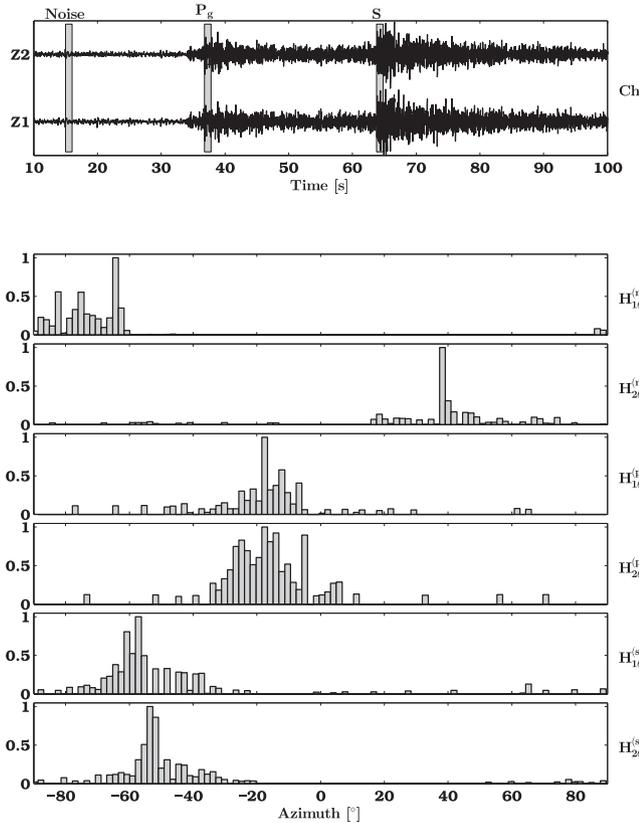


Figure 6. Azimuth histograms for the earthquake of Fig. 5, from seismograms at the two shallowest sensors of the borehole array. Top plot shows raw Z component data with histogram windows highlighted. Second and third plots show histograms of noise; fourth and fifth plots, P_g ; last two plots, S . Superscripts in plot labels indicate phase: (p) for P_g , (s) for S , (n) for noise. Subscripts indicate sensor number. Times are relative to event origin.

seismic phases could be identifiable from polarization alone—even for emergent, quasi-continuous signals in noisy environments.

In the rest of this work, we measure similarity using eq. (12) and omit the superscript (α) from $S_\alpha^{(\alpha)}$. When working with real data, D_U , D_L can be obtained using the procedure described in the previous section, with S/N determined from ratios of short- and long-term averages (STA-LTA) of the signal envelope. The advantages of these over absolute values (or squares) of seismogram amplitude are discussed in e.g. Jones & van der Baan (2015) and Earle & Shearer (1994).

4 APPLICATION

In this section, we demonstrate application of the method to two data sets: broad-band seismograms recorded by Canadian National Seismic Network (CNSN) permanent stations, and borehole records from a 12-station array at Hoadley gas field, central Alberta, Canada. The Hoadley array was installed as part of a temporary microseismic monitoring experiment that ran from September 2012 to July 2013; instrumentation details can be found in Appendix A, and a full description of the experiment appears in Eaton *et al.* (2014). Data preprocessing is described in Appendix B and preprocessing parameters for all data are given in Table 3.

Array records of distant earthquakes are a natural test of polarization similarity because epicentral distance is much greater than the distance between any two stations. We expect that a robust similar-

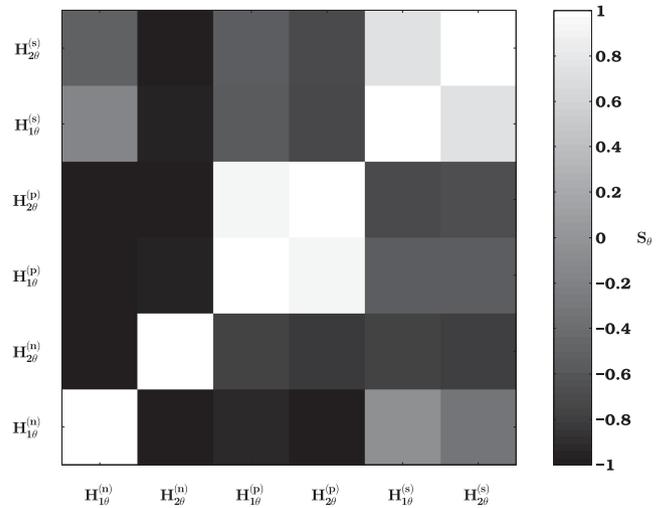


Figure 7. Grey-level image of polarization similarity for the azimuth histograms of Fig. 6. Distance metric and computational details are described in the text. Label superscript and subscripts are as for Fig. 6.

ity measure will produce quantitative results consistent with seismic wave theory, e.g.

- (i) Nearly colocated sensors should have similarly polarized wavefields during phase onsets.
- (ii) Azimuths for P_g and S_g should be dissimilar (low S_θ) at the same sensor.
- (iii) Polarizations of noise, and phase onsets compared with noise, should be neither consistently similar nor consistently dissimilar.

We now test these predictions on two examples of array records. In these examples, all histograms have $N_i = 100$ bins.

Fig. 5 shows a typical example of regional seismicity in western Alberta, Canada: an $M_N = 3.2$ event. The seismograms were recorded by a 12-station borehole array at Hoadley gas field, central Alberta, Canada. Experiment details and data preprocessing are given in Appendix A and Table 3.

We now analyse the polarization similarity of this record. As shorthand we will use $S_{\alpha k}$ to denote the average cross-sensor similarity of attribute α between sensor k and all sensors $k_1 \neq k$.

We begin with the example of azimuth. Fig. 6 shows azimuthal histograms for the two shallowest sensors in the array. The time windows chosen for analysis include the P_g ($t = 35.6$ – 36.6 s) and S_g ($t = 63.9$ – 64.9 s) phase onsets, along with a window of pre-event noise for comparison purposes. Phase onsets are determined by analyst inspection using a location and velocity model provided by Earthquakes Canada and the Alberta Geological Survey, respectively (Caffagni *et al.* 2015).

Fig. 7 shows a greyscale intensity map of azimuthal similarity for the test data. θ_1 , θ_2 are similar for each phase arrival and P_g and S_g are dissimilar. Noise is not coherent between sensors, despite peaks in θ that could indicate polarized noise. We conclude that eq. (12) satisfies the above predictions.

Our picture of the polarization changes little when we use all available station pairs. We will use the variable \bar{S}_α to denote the average of all $K(K - 1)/2$ similarities (across a K -station array) for attribute α . Fig. 9 shows a colour intensity map of $S_{k\alpha}$ for all available geophones in the Hoadley array (designated H01, H02, etc.). Similarity computations use 1 s (100 sample) histogram windows spaced every 0.2 s. Cross-station similarity increases for all

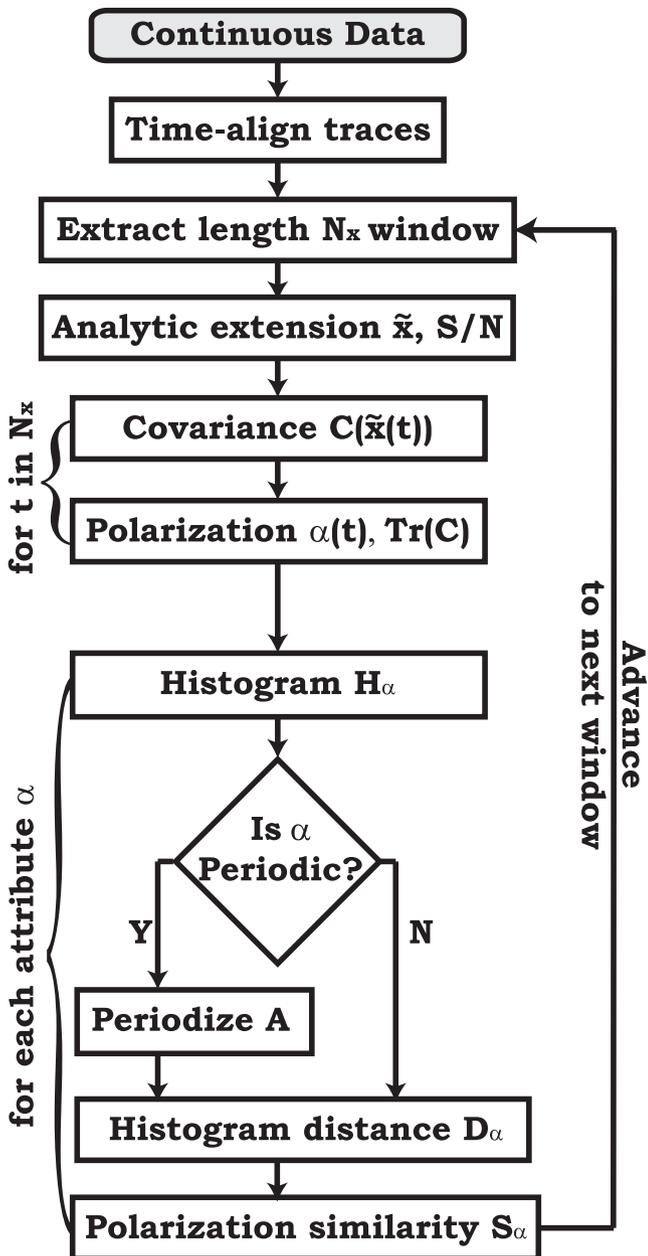


Figure 8. Workflow diagram for determining polarization similarity. \tilde{x} is computed from eq. (1). S/N uses the maximum STA/LTA of the signal envelope in each window (Kanasewich 1981). C is computed from eq. (2), and polarization attributes α use the formulae in Table 1. Histograms use the weighting scheme in eq. (6). D_α uses eq. (8) with $N_i = 100$, $m = 0.5$ and A computed from eq. (10).

attributes at the P_n onset and remains high until the coda onset. Angular polarization is only slightly similar during the P_g phase onset, presumably due to both the emergent arrival and strong scattering in the hydraulic fracture network below the array. ν and ρ are very similar during phase arrivals.

5 DISCUSSION

Fig. 8 describes the full workflow of our method; note that only the last steps are new. One of the great strengths of this approach is that

histogram distance metrics don't assume a unimodal distribution. Thus, even if many signals are present in a single window, and each signal is well-polarized with comparable energy, our similarity measure requires no modification to accurately characterize how polarization changes between stations.

A quantitative measure of polarization similarity has many potential applications. For example, relief plots of polarization similarity allow easy graphical identification of misaligned sensors based on streaks of consistently low or high similarity values, particularly when similarity is unchanged (or decreases, rather than increases) during phase onsets. Consistently dissimilar azimuths ($S_{\theta k} < 0$ everywhere) indicate noisy horizontal components; a decrease in $S_{\theta k}$ during phase arrivals indicates misaligned horizontal components at sensor k . A decrease in $S_{k,\theta}$ during phase arrivals can indicate a rotated vertical component, while consistently low $S_{\theta k}$ indicates a noisy vertical channel. Comparing Fig. 9 with the trace data in Fig. 5 confirms the latter interpretation for sensor H08.

5.1 Effects of station spacing

Noise becomes dissimilarly polarized as station spacing increases, but attenuation and scattering will reduce attribute similarity of incident seismic phases. To investigate how these factors affect our method, we compare records of the 2013 May 24 May $M_w = 8.3$ Sea of Okhotsk earthquake (Ye *et al.* 2013) from two arrays: the Hoadley, AB borehole array, and five permanent CNSN broadband stations in British Columbia and Yukon Territory, Canada (Fig. 10).

Fig. 11 shows a relief image of polarization similarity for (time-aligned) CNSN broad-band data. Polarization histograms are computed in 2 s windows spaced every 0.2 s. Station spacing ranges from 0.3° – 1.8° ; the resultant spread in azimuths is $\sim 3^\circ$, which reduces $S_{\theta k}$ as expected (top plot). The angular polarization measures are dissimilar for the noise but ν , ρ are always extremely similar. Because the oceanic microseism is largely removed by a Butterworth band reject filter, another effect may be responsible for this similarity.

When analysing data from a narrow aperture array, it is normally assumed that polarization is similar for high S/N phase arrivals from a distant source; this assumption underlies the use of array stacking in many seminal polarization papers (e.g. Jurkevics 1988; Bataille & Chiu 1991; Earle 1999). To test how this assumption is borne out by eq. (12), Fig. 12 shows a relief image of $S_{\theta k}$ for the P arrival of the Sea of Okhotsk earthquake at the Hoadley, AB borehole array (red triangle, Fig. 5). Histograms use the same window length and spacing as Fig. 11. The Hoadley geophones have 15–30 m station spacing and a total array aperture of ~ 230 m (Caffagni *et al.* 2015). Thus, the phase onsets should have nearly identical polarizations. The figure shows that these expectations are consistent with values obtained from eq. (12). Incidence angle (ϕ) shows the least similarity; this suggests that either noise affects some vertical component channels (e.g., H08, H10) or that these sensors deviate slightly from vertical. Because H8 and H10 also have the least similar rectilinearity (ρ) during most of the event, the first explanation is more likely.

It is noteworthy that Figs 11 and 12 suggest that either polarized noise is present on both arrays, or that planarity and rectilinearity are not especially diagnostic for identifying teleseismic phase arrivals. This issue partly arises because our adaptive similarity measure was created using Monte-Carlo simulations of ν , $\rho \in [0, 1]$, but the conditions required for ν , $\rho \sim 0$ are difficult to achieve in real

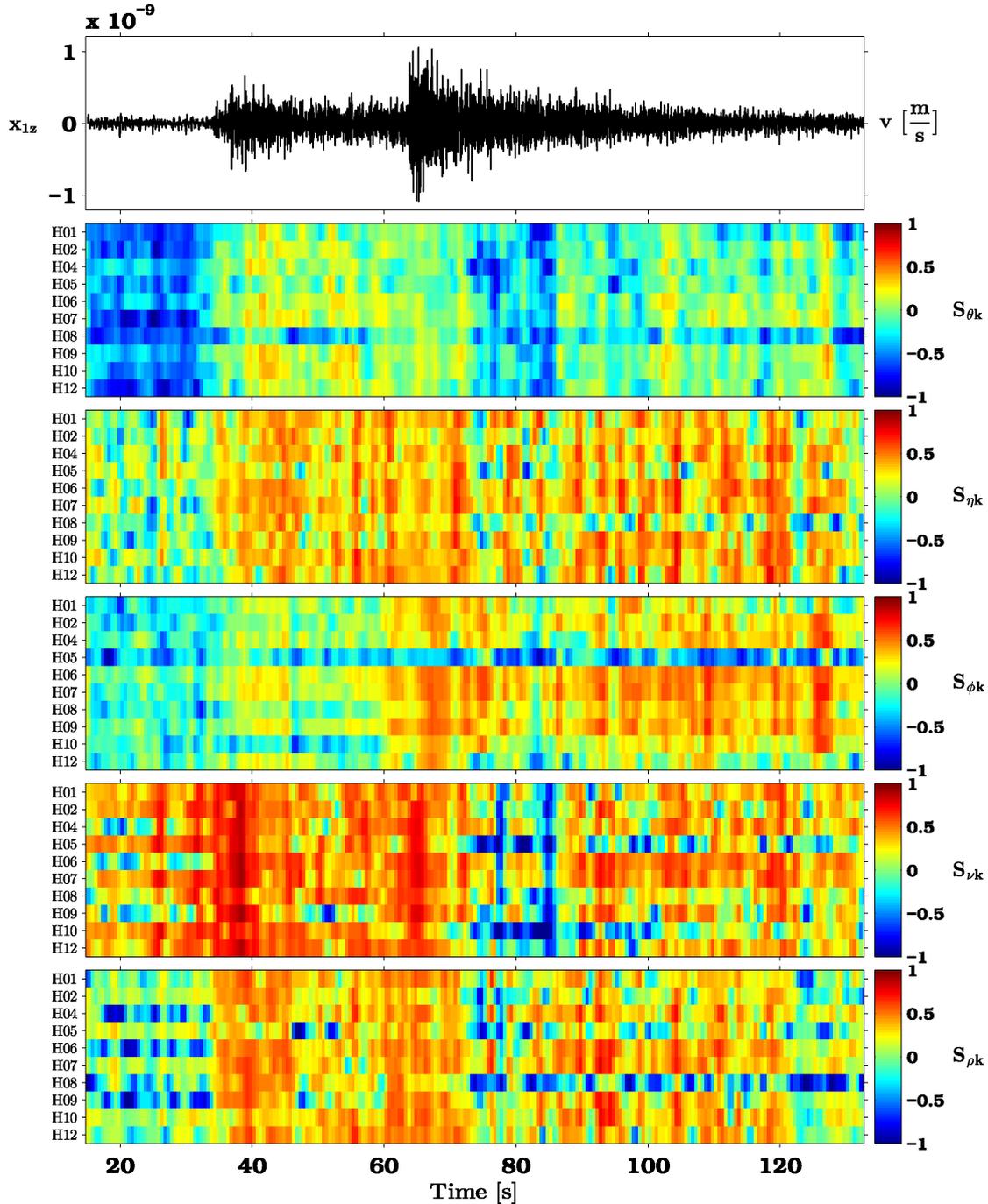


Figure 9. Colour intensity map of cross-station polarization similarity ($S_{\alpha k}$) for the event and array of Fig. 5. Second subscript in each label indicates polarization attribute (Table 1). Top plot shows z -component data from the shallowest borehole geophone ($z = -1606$ m a.s.l.). Times are relative to earthquake origin.

seismic data. From their definitions in Jurkevics (1988), low values of rectilinearity and planarity require 2+ out-of-phase wavefields with comparable amplitudes; for the data and short averaging windows in this work, these conditions are only obtainable during seismic codas (e.g. Fig. 9). Exotic scenarios can achieve these conditions more reliably: for example, at persistently active basaltic volcanoes, multiple volcanic tremor sources can combine to produce low ν & ρ values in frequency bands where different tremor sources have comparable amplitudes (Jones *et al.* 2012a,b). Yet per-

sistent quasi-continuous tremor with multiple active sources is only known to exist at a few volcanoes worldwide: examples include Erta 'Ale, Ethiopia (Jones *et al.* 2012a,b), Etna, Italy (Acernese *et al.* 2004), and Marum (Ambrym), Vanuatu (Carniel *et al.* 2003).

Because $\nu, \rho \in [0, 1]$ can be seen in field data, our empirical ranges for S_ρ, S_ν , are appropriate. However, in specific cases where planarity and rectilinearity are of greater interest and coda waves are not being studied, it may be more diagnostic to rederive their similarity measures using Monte-Carlo simulations over a narrower

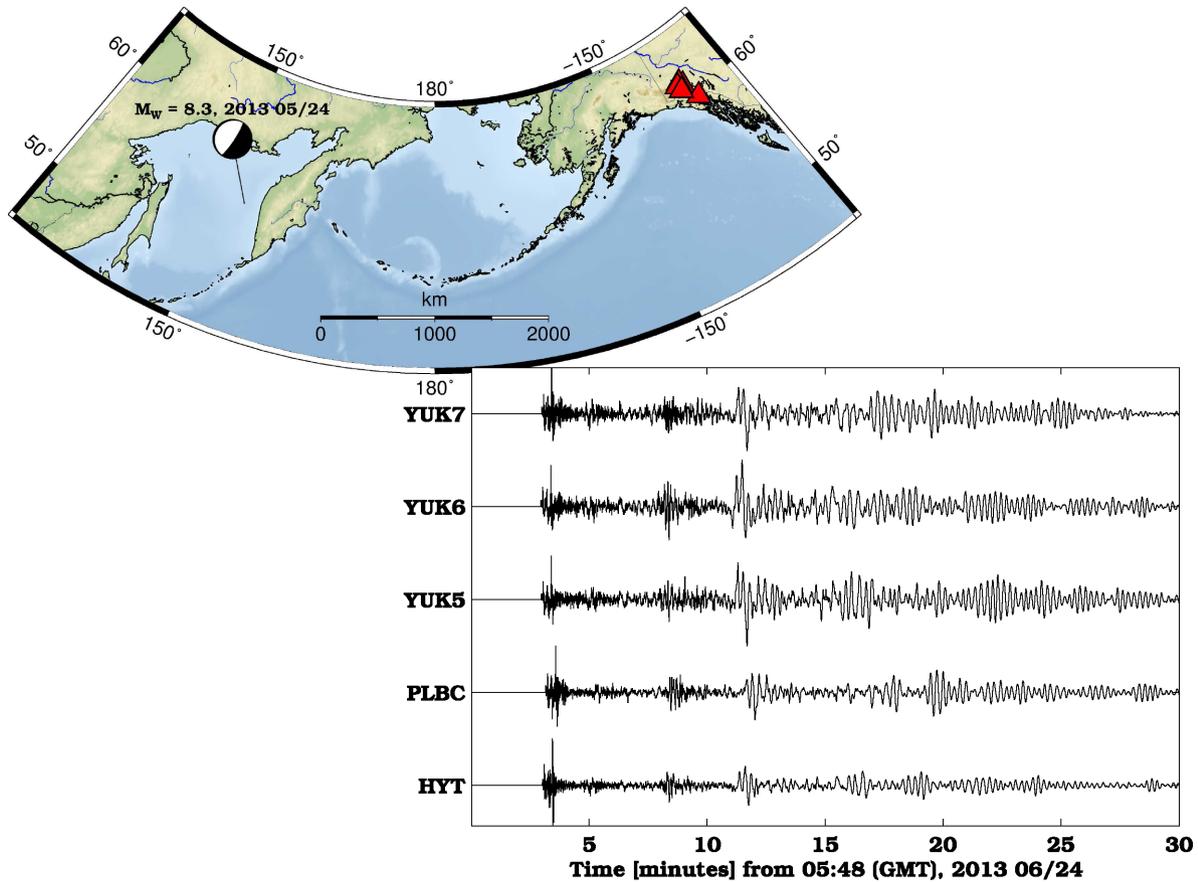


Figure 10. Source–receiver geometry and time-aligned vertical component seismograms from the $M_w = 8.3$ Sea of Okhotsk earthquake, 2013 May 24 (UTC), recorded by Canadian National Seismic Network (CNSN) permanent broad-band sensors (red triangles).

attribute range; we caution, however, that this will result in S_ρ , $S_v \leq -1$ during seismic codas. We used the range $[0, 1]$ mainly to remain consistent with the other attributes studied in this work.

5.2 Time–frequency polarization similarity

Time–frequency polarization analysis is commonly used to isolate noise and unwanted secondary signals (e.g. Lilly & Park 1995; Anant & Dowla 1997; Gao *et al.* 1999; Kulesh *et al.* 2007; D’Auria *et al.* 2010; Jones *et al.* 2012a). However, caution must be exercised, as most time–frequency transforms implicitly average many data points to create each coefficient. For example, each coefficient in the continuous wavelet transform (CWT) at any scale L has a time-domain ‘cone of influence’ $C_w \propto LB$, where B is the mother wavelet’s filter width (Mallat 1989; Daubechies 1992; Kumar & Foufoula-Georgiou 1997). Similarly, the discrete wavelet transform (DWT) at any level j measures changes in averages of X_t on scale lengths $L = 2^j$, giving each wavelet coefficient W_{jt} a cone of influence $C \propto 2^j F$ (Percival & Walden 2000, chap. 4). Thus the polarization of wavelet coefficients is never truly instantaneous; this implicit averaging reduces the energy of seismic phase arrivals relative to that of background noise. Thus, eq. (12) is well-suited to time–frequency extension, but a static similarity measure like eq. (11) would perform poorly in the wavelet domain.

We now demonstrate time–frequency extension of polarization similarity using the undecimated ‘maximal overlap’ DWT

(MODWT) of Walden & Cristan (1998). Computing polarization from a complex CWT produces equivalent information to computing polarization from the analytic extension of the MODWT at scales $j = 2^L$ (for scale parameter L , $j = 1, 2, \dots, J$). This follows from the fact that the MODWT can be formed from the CWT taken at these scales; exact equivalence is demonstrated in Percival & Walden (2000) and Walden & Cristan (1998, chap. 5). Due to the choice of mother wavelet, very minor differences can arise between this approach and CWT-based techniques (e.g. Gao *et al.* 1999; Kulesh *et al.* 2007); the CWT also allows more freedom in choice of scales if perfect reconstruction is not required. In this work, we use a length 16 ‘least asymmetric’ mother wavelet because of its excellent frequency localization properties (Jones *et al.* 2012a, fig. 1).

Fig. 6 suggests that polarized noise may affect some sensors in the geophone array; this is somewhat surprising considering the sample event postdates peak resource extraction by several months. Time–frequency analysis is a natural way to constrain the potential source(s). For example, sensor installation errors (e.g. poor geophone clamping) are a well-documented noise source in borehole arrays (St-Onge & Eaton 2011; St-Onge *et al.* 2013). If polarizations are similar in part of the frequency spectrum before the P onset in the sample data, then noise generated by individual sensors is unlikely to be the cause.

Figs 13 and 14 show the resulting time–frequency polarization similarity for the earthquakes of Figs 5 and 10, respectively. Note that data are detrended for this procedure, but not bandpass filtered;

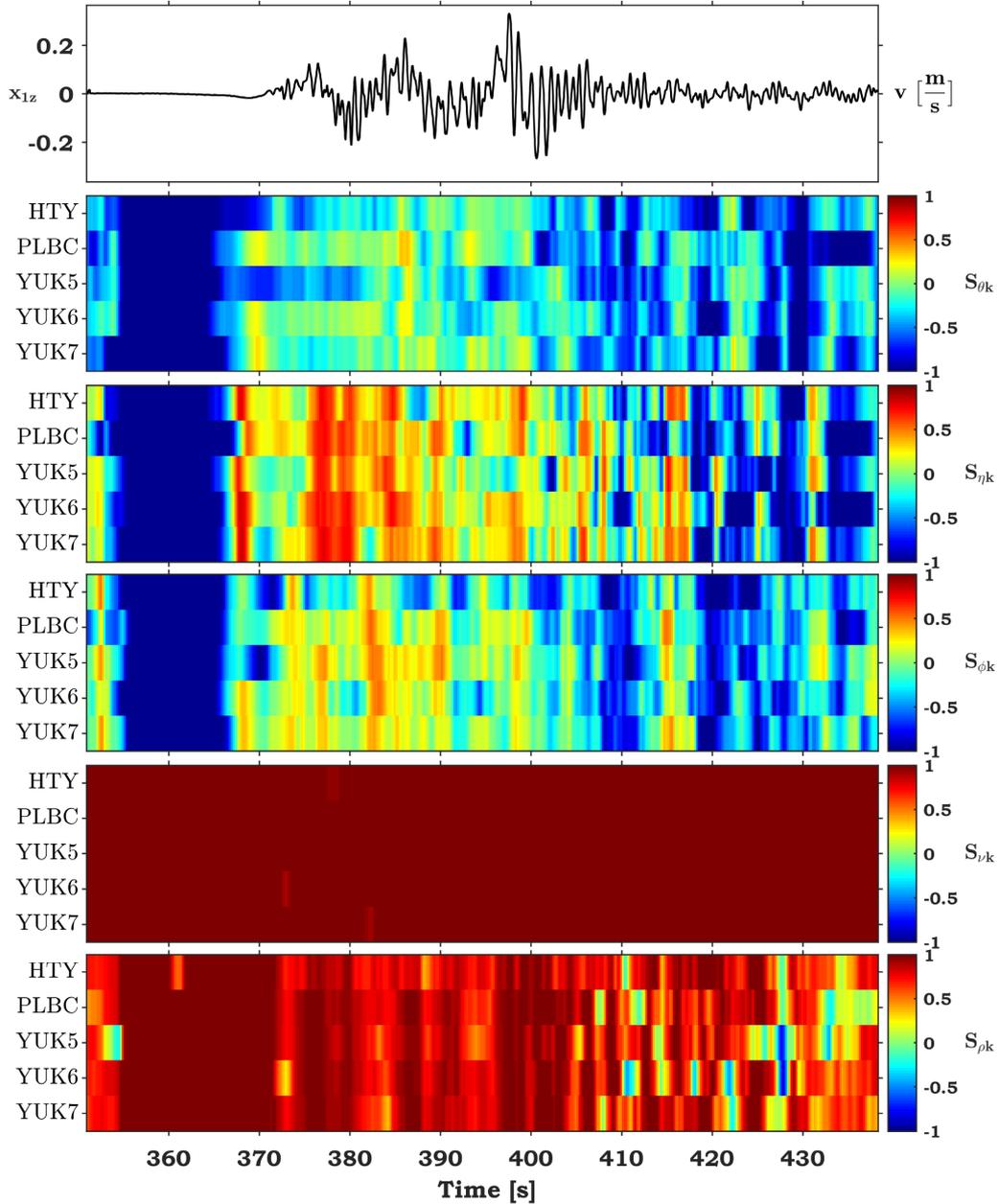


Figure 11. Relief plot of $S_{\alpha k}$ for the P arrival of the 2013 $M_w = 8.3$ Sea of Okhotsk earthquake using trace data from CNSN permanent broad-band stations (Fig. 10). Top plot shows z -component data from station HTY as a reference. Vertical axes on relief images are labelled with sensor names. To generate this figure, in addition to the preprocessing in Table 3, seismograms were band reject filtered with a two-pole Butterworth filter (corners 0.2 and 0.8 Hz) to remove the oceanic microseism.

however, polarization similarity is not shown for the scaling coefficients because the $f_c = 15$ Hz geophones are insensitive to very low frequencies (Caffagni *et al.* 2015). Histogram windows are length 1 s for each event, computed every 0.2 s.

It is immediately clear that the noise is similarly polarized across the entire array at frequencies $f \geq 5$ Hz. Because P_g is emergent, and the upgoing waves are presumably scattered by the heterogeneous structure of the hydraulic fracture network, $\bar{S}_{\theta k}$ and $\bar{S}_{\phi k}$ decrease at the P_g onset in Fig. 14. We therefore conclude that the similar polarizations cannot arise from instrument self-noise or other localized sensor noise. Anthropogenic noise is also unlikely because the Sea of Okhotsk event occurred at 11:44 p.m. local time

(UTC -6). It therefore appears that a persistent, high-frequency ($f \geq 5$ Hz) seismic source was active near the Hoadley, AB array at both times.

Coherent noise and tremor-like signals were previously seen in hydraulic fracture monitoring by Tary *et al.* (2014), with the former ascribed to mechanical pumping. However, spectrograms (e.g. Fig. 5) lack the characteristic narrowband resonances of the latter. Given that these records postdate hydraulic fracture treatment (2012 September 17–18) and peak resource extraction (December 2012–January 2013), the noise source could be mechanical pumping from another hydraulic fracture. However, the possibility of a tremor-like source arising from fluid propagation in the hydraulic

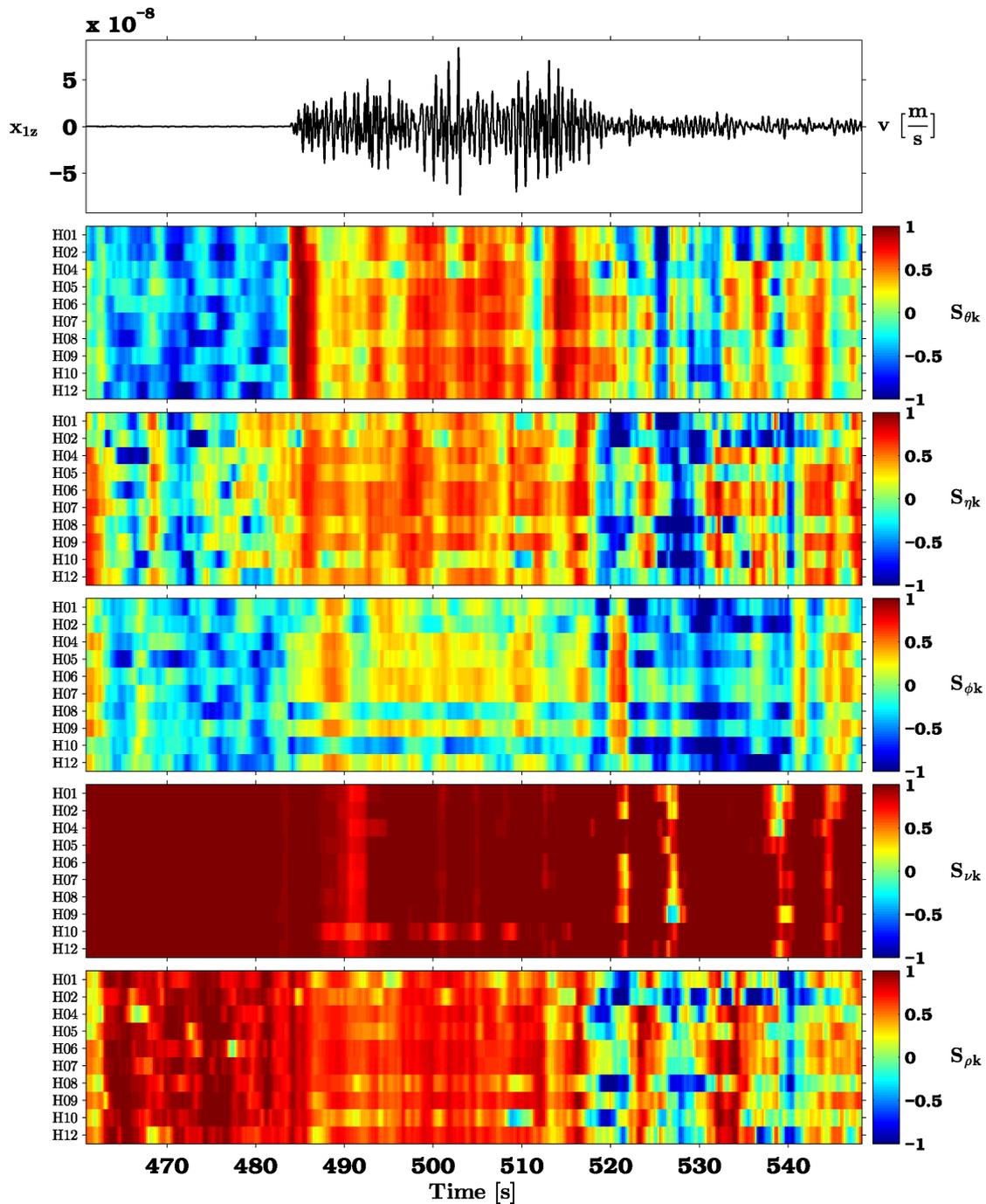


Figure 12. Relief plot of $S_{\alpha k}$ for the P arrival of the 2013 $M_w = 8.3$ Sea of Okhotsk earthquake at the Hoadley, AB array (Fig. 5). Top plot shows z -component data from the shallowest geophone in the array. Vertical axes on relief images are labelled with sensor names.

fracture network, with a different source process than that described in Tary *et al.* (2014), cannot be ruled out. These possibilities are speculative, however. As of this writing, only a few works have focused on low-frequency signals recorded during hydraulic fractures (e.g. Tary & van der Baan 2012; Das & Zoback 2013a,b; Tary *et al.* 2014); none of their reported signal characteristics match the pre-arrival noise in Figs 9 and 12–14. Thus, constraining possible noise sources and their relationships (if any) with hydraulic fracture treatment will be the subject of future work.

6 CONCLUSIONS

In this paper, we introduce a robust method for quantitative comparison of seismic polarization attributes. Using overlapping windows of attribute histograms, we can study how polarization similarity changes before and during phase arrivals. Because our method determines similarity in a way that accounts for the varying signal-to-noise ratios of real data, this technique could provide an effective means of identifying seismic phases in noisy environments. Potential uses include diagnosing site effects, identifying reflected and

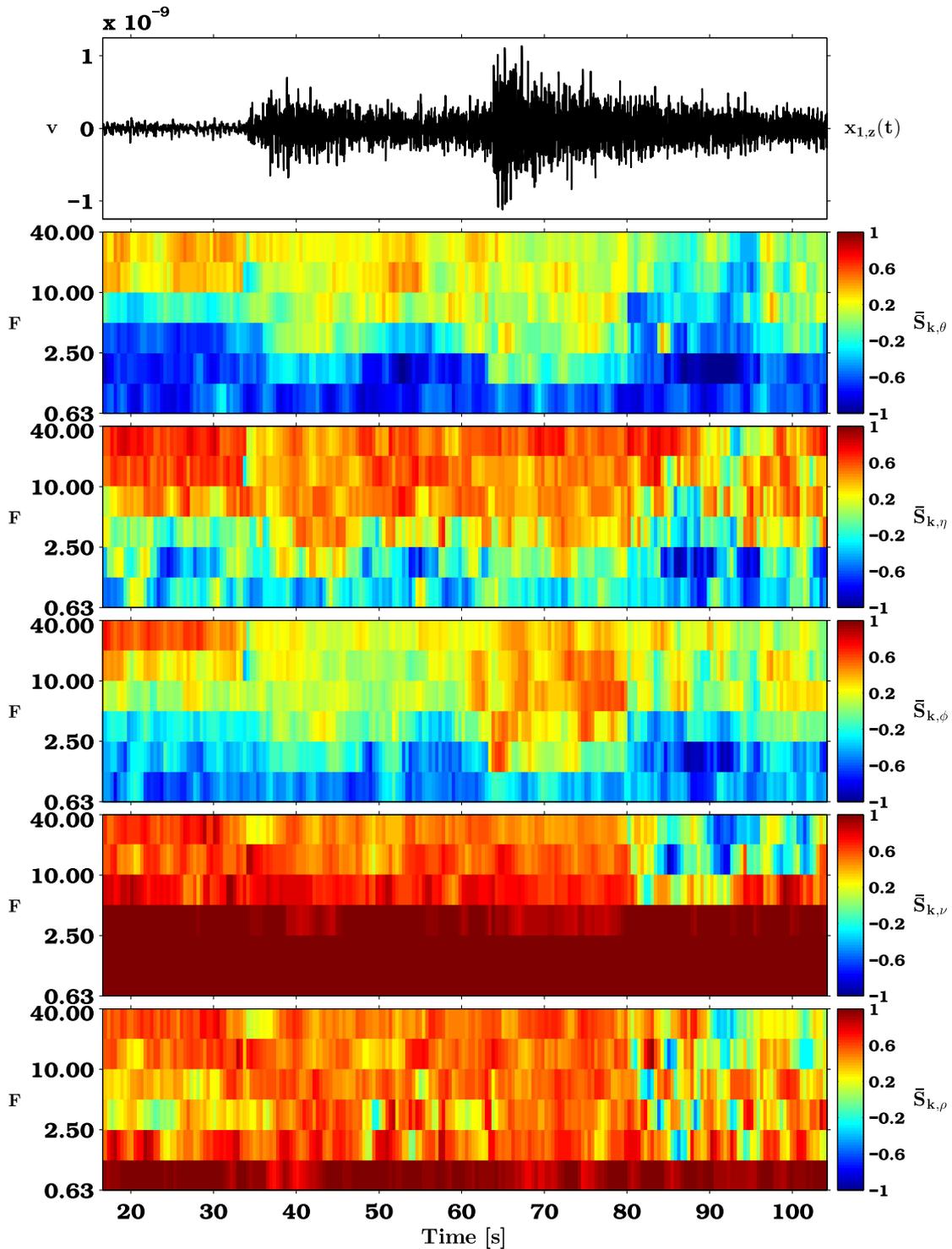


Figure 13. Time–frequency decomposition of average polarization similarity for each attribute, using the earthquake and array of Fig. 5 and MODWT detail coefficients $\tilde{W}_j^{(D)}(j = 1 : 1 : 5)$. Frequency (y-axis) uses an octave scale. Top plot shows z-component data from geophone H01.

converted phases, and detecting noisy or misaligned sensors without the need for additional perforation shots. In the latter case, detecting relative misalignment of sensors within an array is possible even when the absolute sensor orientations are unknown. Our technique is easily implemented as an extension of exist-

ing polarization analysis routines. Applying our method to noise recorded in a borehole at Hoadley, AB demonstrates identification of sensor noise; time–frequency analysis suggests a seismic source persisted near the monitoring array several months after peak resource extraction.

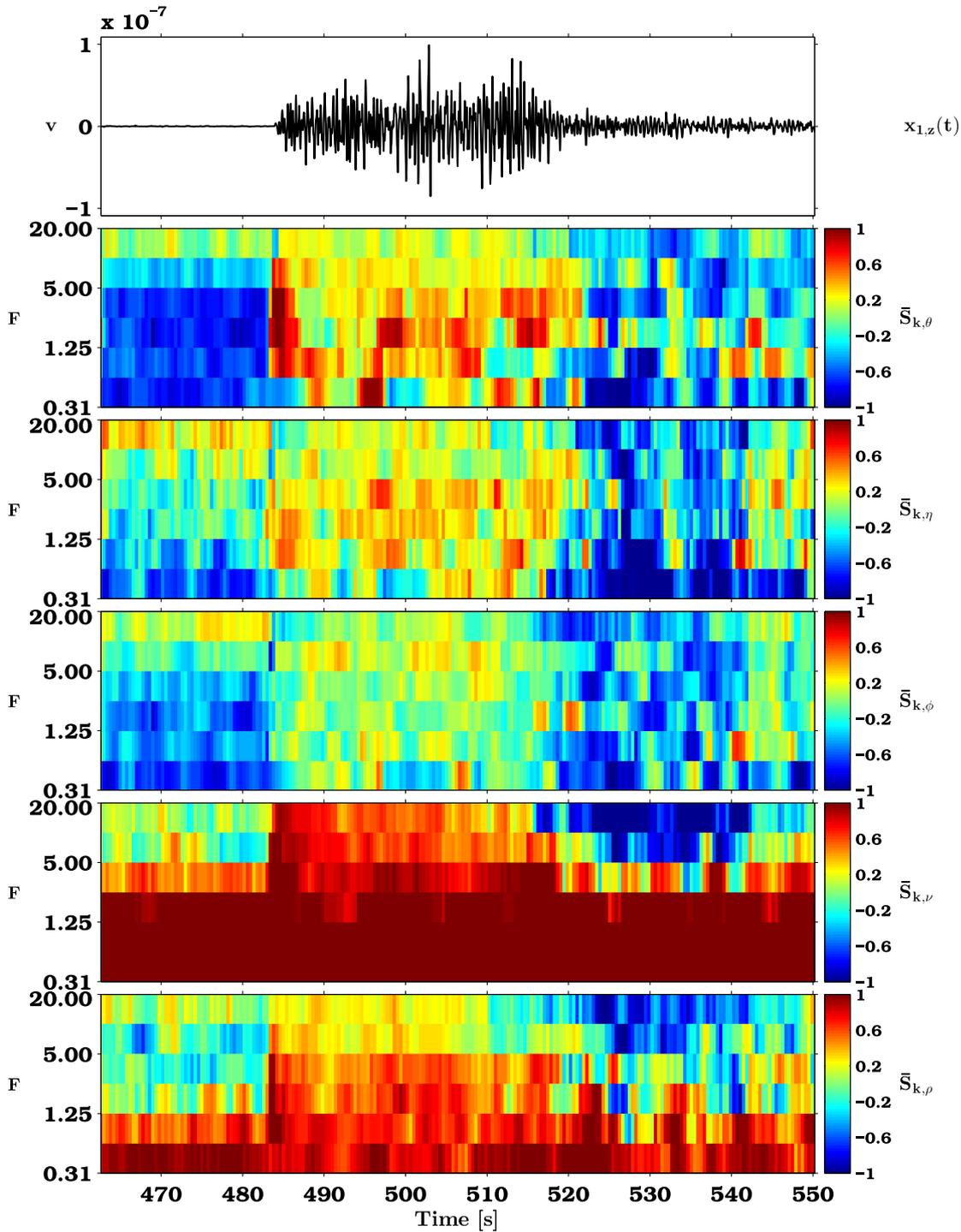


Figure 14. Time–frequency decomposition of average polarization similarity for each attribute of the 2013 Sea of Okhotsk earthquake (Fig. 10), using data from the Hoadley, AB geophone array (Fig. 5) and MODWT detail coefficients $\tilde{W}_j^{(D)}$ ($j = 1 : 1 : 5$). Frequency (y -axis) uses an octave scale. Top plot shows z -component data from geophone H01.

ACKNOWLEDGEMENTS

We thank the Microseismic Industry Consortium (jointly operated by the University of Alberta, University of Calgary, and numerous industry sponsors) for data and financial support, and Conoco Phillips Canada and the Natural Sciences and Engineering Research Council for financial support for the Hoadley experiment. We thank the National WaveForm Archive (NWA) of Natural Resources

Canada for providing the Yukon and British Columbia station data. JJ extends special thanks to J. Vidale (University of Washington) and M. van der Baan (University of Alberta) for insights and discussion of polarization analysis. We extend special thanks to our two anonymous reviewers and Dr Joshua Carmichael (Los Alamos National Laboratory, USA) for their reviews, which helped to clarify and improve the manuscript.

REFERENCES

- Acernese, F., Ciaramella, A., De Martino, S., Falanga, M., Godano, C. & Tagliaferri, R., 2004. Polarization analysis of the independent components of low frequency events at Stromboli volcano (Eolian Islands, Italy), *J. Volc. Geotherm. Res.*, **137**(1), 153–168.
- Anant, K.S. & Dowla, F.U., 1997. Wavelet transform methods for phase identification in three-component seismograms, *Bull. seism. Soc. Am.*, **87**, 1598–1612.
- Bataille, K. & Chiu, J.M., 1991. Polarization analysis of high-frequency three-component seismic data, *Bull. seism. Soc. Am.*, **81**, 622–642.
- Caffagni, E., Eaton, D., van der Baan, M. & Jones, J.P., 2015. Regional seismicity: a potential pitfall for identification of long-period long-duration events, *Geophysics*, **80**(1), A1–A5.
- Carniel, R., Di Cecca, M. & Rouland, D., 2003. Ambrym, Vanuatu (July–August 2000): spectral and dynamical transitions on the hours-to-days timescale, *J. Volc. Geotherm. Res.*, **128**(1), 1–13.
- Das, I. & Zoback, M.D., 2013a. Long-period, long-duration seismic events during hydraulic stimulation of shale and tight-gas reservoirs – Part 1: Waveform characteristics, *Geophysics*, **78**(6), KS97–KS108.
- Das, I. & Zoback, M.D., 2013b. Long-period long-duration seismic events during hydraulic stimulation of shale and tight gas reservoirs – Part 2: Location and mechanisms, *Geophysics*, **78**(6), KS109–KS117.
- Daubechies, I., 1992. Ten lectures on wavelets, in *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial And Applied Mathematics, 377 pp.
- D'Auria, L., Giudicepietro, F., Martini, M., Orazi, M., Peluso, R. & Scarpato, G., 2010. Polarization Analysis in the Discrete Wavelet Domain: An Application to Volcano Seismology, *Bull. seism. Soc. Am.*, **100**, 670–683.
- De Meersman, K., van der Baan, M. & Kendall, J.-M., 2006. Signal extraction and automated polarization analysis of multicomponent array data, *Bull. seism. Soc. Am.*, **96**, 2415–2430.
- Earle, P. & Shearer, P., 1994. Characterization of global seismograms using an automatic picking algorithm, *Bull. seism. Soc. Am.*, **84**(2), 366–376.
- Earle, P.S., 1999. Polarization of the Earth's teleseismic wavefield, *Geophys. J. Int.*, **139**(1), 1–8.
- Eaton, D., Caffagni, E., Rafiq, A., van der Baan, M., Roche, V. & Matthews, L., 2014. Passive seismic monitoring and integrated geomechanical analysis of a tight-sand reservoir during hydraulic-fracture treatment, flowback and production, in *Proceedings of the 2nd Unconventional Resources Technology Conference (URTEC)*, doi:10.15530/urtec-2014-1929223.
- Gao, J., Dong, X., Wang, W.B., Li, Y. & Pan, C., 1999. Instantaneous parameters extraction via wavelet transform, *IEEE Trans. Geosci. Remote Sens.*, **37**(2), 867–870.
- Hafner, J., Sawhney, H.S., Equitz, W., Flickner, M. & Niblack, W., 1995. Efficient color histogram indexing for quadratic form distance functions, *IEEE Trans. Pattern Anal. Mach. Intell.*, **17**(7), 729–736.
- Jones, J.P. & van der Baan, M., 2015. Adaptive STA-LTA with outlier statistics, *Bull. seism. Soc. Am.*, **105**(3), 1606–1618.
- Jones, J.P., Carniel, R. & Malone, S.D., 2012a. Subband decomposition of continuous volcanic tremor, *J. Volc. Geotherm. Res.*, **213**, 98–115.
- Jones, J.P., Carniel, R. & Malone, S.D., 2012b. Decomposition, location, and persistence of seismic signals recovered from continuous tremor at Erta 'Ale, Ethiopia, *J. Volc. Geotherm. Res.*, **213**, 116–129.
- Jurkevics, A., 1988. Polarization analysis of three-component array data, *Bull. seism. Soc. Am.*, **78**, 1725–1743.
- Kanasewich, E.R., 1981. *Time Sequence Analysis in Geophysics*, University of Alberta Press.
- Kulesh, M., Diallo, M., Holschneider, M., Kurennaya, K., Kruger, F., Ohrnberger, M. & Scherbaum, F., 2007. Polarization analysis in the wavelet domain based on adaptive covariance method, *Geophys. J. Int.*, **170**, 667–678.
- Kumar, P. & Foufloula-Georgiou, E., 1997. Wavelet analysis for geophysical applications, *Rev. Geophys.*, **35**(4), 385–412.
- Lilly, J.M. & Park, J., 1995. Multiwavelet spectral and polarization analyses of seismic records, *Geophys. J. Int.*, **122**, 1001–1021.
- Ling, H. & Okada, K., 2007. An efficient earth mover's distance algorithm for robust histogram comparison, *IEEE Trans. Pattern Anal. Mach. Intell.*, **29**(5), 840–853.
- Mallat, S., 1989. Multifrequency channel decompositions of images and wavelet models, *IEEE Trans. Acoust. Speech Signal Process.*, **37**, 2091–2110.
- Monge, G., 1781. *Mémoire sur la théorie des déblais et des remblais*, De l'Imprimerie Royale.
- Montalbetti, J.R. & Kanasewich, E.R., 1970. Enhancement of teleseismic body phases with a polarization filter, *Geophys. J. R. astr. Soc.*, **21**, 119–129.
- Munkres, J.R., 1975. *Topology: A First Course*, Prentice-Hall.
- Park, J., Vernon, F.L. III & Lindberg, C.R., 1987. Frequency dependent polarization analysis of high-frequency seismograms, *J. geophys. Res.*, **92**, 12 664–12 674.
- Pele, O. & Werman, M., 2010. The quadratic-chi histogram distance family, in *Computer Vision—ECCV 2010*, pp. 749–762, eds Daniilidis, K., Maragos, P. & Paragios, N., Springer.
- Percival, D.B. & Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*, Cambridge Univ. Press.
- Pinnegar, C.R., 2006. Polarization analysis and polarization filtering of three-component signals with the time-frequency S transform, *Geophys. J. Int.*, **165**(2), 596–606.
- Rene, R.M., Fitter, J.L., Forsyth, P.M., Kim, K.Y., Murray, D.J., Walters, J.K. & Westerman, J.D., 1986. Multicomponent seismic studies using complex trace analysis, *Geophysics*, **51**(6), 1235–1251.
- Rubner, Y., Tomasi, C. & Guibas, L.J., 2000. The earth mover's distance as a metric for image retrieval, *Int. J. Comput. Vis.*, **40**(2), 99–121.
- Snedecor, G.W. & Cochran, W.G., 1967. *Statistical Methods*, 6th ed., Iowa State University Press.
- St-Onge, A. & Eaton, D.W., 2011. Noise examples from two microseismic datasets, *CSEG Recorder*, **36**(10), 46–49.
- St-Onge, A., Eaton, D.W. & Pidlisecky, A., 2013. Borehole vibration response to hydraulic fracture pressure, paper presented at CSEG Geoconvention, Calgary.
- Tary, J.B. & Van der Baan, M., 2012. Potential use of resonance frequencies in microseismic interpretation, *Leading Edge*, **31**, 1330–1337.
- Tary, J.B., Baan, M. & Eaton, D.W., 2014. Interpretation of resonance frequencies recorded during hydraulic fracturing treatments, *J. geophys. Res.*, **119**(2), 1295–1315.
- VanDecar, J.C. & Crosson, R., 1990. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. seism. Soc. Am.*, **80**(1), 150–169.
- Vidale, J.E., 1986. Complex polarization analysis of particle motion, *Bull. seism. Soc. Am.*, **76**, 1393–1405.
- Walden, A.D. & Cristan, A.C., 1998. The phase-corrected undecimated discrete wavelet packet transform and the recurrence of high latitude interplanetary shock waves, *Proc. R. Soc. A*, **454**, 2243–2266.
- Ye, L., Lay, T., Kanamori, H. & Koper, K.D., 2013. Energy release of the 2013 Mw 8.3 Sea of Okhotsk earthquake and deep slab stress heterogeneity, *Science*, **341**(6152), 1380–1384.

APPENDIX A: DATA AND RESOURCES

Two data sets were used in this paper.

Data for the $M_w = 8.3$ Sea of Okhotsk earthquake were obtained by anonymous FTP using the CNSN AutoDRM online request system in September 2014. All CNSN stations used in this paper are CMG-3T broad-band sensors ($f_c = 1/120$ Hz, $h_d = 0.707$) sampling continuously at $f_s = 100$ Hz. The stations occupy permanent

monitoring sites in southern Yukon Territory and northern British Columbia, Canada. These data may be requested directly from the CNSN at: http://www.earthquakescanada.nrcan.gc.ca/stndon/AutoDRM/autodrm_req-eng.php.

Regional earthquake data were provided by the Microseismic Research Consortium, a joint venture of the University of Alberta, University of Calgary, and numerous industry sponsors. In 2012–2013, the consortium deployed a vertical borehole array of three-component passive triaxial geophones ($f_c = 15$ Hz, $h_d = 1.0$) to monitor an open-hole multistage hydraulic fracture treatment at Hoadley gas field, central Alberta, Canada. The experiment recorded several months of continuous data at $f_s = 4000$ Hz. Station spacing ranged from ~ 15 to 30 m across the array, with total vertical aperture ~ 230 m. Two sensors (H03 & H11) were excluded from the analysis due to sensor noise and potential timing issues. These data may be requested by contacting Prof David W. Eaton.

All data analysis programs were written in Matlab R2014a (8.03). They are available upon request from Dr Joshua Jones.

APPENDIX B: PREPROCESSING

All data were converted to ASCII format, detrended and downsampled from their original sampling rates, then filtered with a two-pole Butterworth bandpass filter. Downsampling and filtering parameters are given in Table 3; all data were corrected to $h_d = 0.707$ at the new critical frequency.

For the borehole array data, instrument response was flattened to $f_c = 3$ Hz for the regional event and $f_c = 0.2$ Hz for the teleseism. Data were then rotated to the [Z N E] coordinate system of eqs (2) and (3) using empirical corrections to sensor azimuths determined from high S/N P arrivals of local and teleseismic earthquakes, including the $M_w = 8.3$ Sea of Okhotsk event.

During polarization analysis all traces from both data sets were time-aligned using optimal shifts ξ_k for Z -component data, as determined via the least-squares waveform cross-correlation technique of Vandecar & Crosson (1990). We imposed the additional constraint $\max(|\xi_k|) = 0.1$ s on the borehole array data to prevent cycle skipping errors.