Signal extraction and automated polarisation analysis of multi-component array data

K. De Meersman

School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK

(Now at: Veritas GeoServices, 715 Fifth Avenue SW, Suite 2200, Calgary,
Alberta, Canada, T2P 5A2, email: kristof_demeersman@veritasdgc.com)

M. van der Baan

School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK

J-M. Kendall

School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK

(Now at: Department of Earth Sciences, University of Bristol, Wills Memorial Building,
Queen's Road Bristol, BS8 1RJ, UK)

May 29, 2006

1 abstract

We present a method to extract seismic signals from three-component array data and estimate their polarisation properties at each station. The technique is based on a Singular Value Decomposition (SVD) of the complex 3C analytic signal and applies to linearly as well as elliptically polarised seismic phases. To increase accuracy we simultaneously analyse data from different stations and apply a noise weighting based on pre-arrival data. For polarisation analysis, an automated routine is also included. It selects the data-window with best signal-to-noise ratio to obtain the polarisation from. A linearity measure and a confidence interval accompany the polarisation estimate at each station in the array. We test our technique for automated polarisation analysis on synthetic P-wave data and compare results to those from other methods. A microseismic dataset from the North Sea provides a unique opportunity to statistically compare previous and independently obtained P-wave polarisations with those provided with the automated technique presented here. We conclude that, for P-wave polarisation analysis, our method is robust and significantly more accurate than conventional, mainly manual methods. This is especially so on data with polarised and correlating background noise. It is also faster and provides meaningful quality estimates.

Keywords: multi-component, array, polarisation analysis, signal separation, SVD

2 Introduction

Signal extraction and polarisation analysis forms an important part of the processing and interpretation of multicomponent seismic data. There are many examples in seismology where polarisation analysis helps with the identification and classification of different seismic phases (Earle, 1999; Flinn, 1965; Mao and Gubbins, 1995; Vidale, 1986). Knowing the polarisation properties of seismic data also helps with the design of so-called polarisation filters. These extract or attenuate signals with specific polarisation properties (Bataille and Chiu, 1991; de Franco and Musacchio, 2001; Du et al., 2000; Montalbetti and Kanasewich, 1970; Reading et al., 2000). Polarisation analysis is also central to the measurement of seismic anisotropy on P-waves (Bear et al., 1999; Schulte-Pelkum et al., 2001) and S-waves (Silver and Chan, 1991; Teanby et al., 2004). Finally, P-wave polarisations can provide estimates of the source azimuth and thereby help constrain earthquake locations (Dyer et al., 1999; Maxwell et al., 1998).

Singular Value Decomposition (SVD), or Eigen Value Decomposition (EVD), are generally used to estimate polarisation properties and extract signal from data. Both techniques are closely related matrix operations. SVD applies to rectangular matrices, where EVD applies to square matrices. In early applications the 3C seismograms would be stored in a N by 3 matrix, with N the number of samples (Flinn, 1965; Montalbetti and Kanasewich, 1970). This time-domain approach suffers from stability problems in short time windows around zero crossings. A further restriction is that it can only deal with linearly polarised phases. Replacing the seismic data by their Fourier transforms, wavelet transforms, or analytic signals allows for the analysis of phases with linear and elliptical polarisations. Fourier and wavelet methods are best to deal with interfering seismic phases with different frequency content (Du et al., 2000; Samson, 1973). Analytic signals permit the use of short analysis windows (Vidale, 1986). This is ideal when dealing with multiple short-duration arrivals that are closely separated in time and have similar frequency content.

Various authors have stacked information from different 3C stations within an array (Bataille and Chiu, 1991; Earle, 1999; Jurkevics, 1988). This improves signal and polarisation estimates since it increases the signal-to-noise ratio. The method is nevertheless restricted to (small) arrays over which signal and polarisation properties remain constant. On the other hand it is possible to combine the N data-samples from k 3C stations into one large N by $3k$ data matrix. SVD and EVD equally apply to these large data matrices. This approach can handle polarisation variations within the array, and also benefits from an increased signal-to-noise ratio (Bear et al., 1999).

EVD and SVD usually rely on the assumption that the background noise is random and isotropic. A biased signal and polarisation estimate is obtained in cases where these assumptions break down. Noise-weighted EVD and SVD can counter the negative effects of polarised background noise (Du et al., 2000; Samson, 1983a,b).

We present an automated time-domain, weighted 3C array technique for robust polarisation estimation. SVD and EVD are central to our approach and following Vidale (1986) we use the analytic signal. We have combined this with multi-station array analysis and noise weighting as per Bear et al. (1999) and Samson (1983a). The technique is applicable to array and repeat source data that are contaminated with coherent and polarised noise. The analysis window is automatically and iteratively optimised by identifying those samples with large deviation from the measured polarisation and excluding them from the improved analysis window in the next iteration.

We will start with the 3C approach of Flinn (1965) and then gradually add the components that make up our method for signal extraction and polarisation analysis. The automated polarisation analysis is introduced next. We then apply the technique on both a synthetic and real P-wave dataset and discuss the results. The real data come from a microseismic experiment in the North Sea Valhall oilfield and conducted in 1998. The examples cover three main issues: (1) estimation and correction for relative arrival-time differences between the different 3C datasets, (2) the derivation of polarisation uncertainty estimates and other quality control parameters, (3) evaluation of accuracy through comparisons with the exact polarisation estimates for the synthetic example and manually obtained estimates for the real dataset.

3 Singular Value Decomposition and Eigen Value Decomposition

3.1 General background

Given a three-component (3C) seismic dataset with the seismograms $n(t)$, $e(t)$ and $z(t)$ recorded in the North, East, and vertical directions, we can construct a N *by* 3 data matrix $\mathbf{D}$ in which each column represents a seismogram over the time-window $t \in [T_1, T_N]$, or $\mathbf{D} = [n(t), e(t), z(t)]$. Singular Value Decomposition (SVD) and Eigen Value Decomposition (EVD) are popular techniques to find the waveform $\mathbf{u}_S$, magnitude σ_S, and polarisation $\mathbf{v}_S$ of the signal that is present in the data $\mathbf{D}$. The SVD of the data in $\mathbf{D}$ is given by

$$\mathbf{D} = \mathbf{U}\Lambda\mathbf{V}^T = \sum_{i=1}^{3} \sigma_i \mathbf{u}_i \mathbf{v}_i^T. \tag{1}$$

where $\mathbf{D}$ is the product of the N *by* 3 matrix $\mathbf{U}$, the 3 *by* 3 diagonal matrix Λ, and the transpose (T) of the 3 *by* 3 matrix $\mathbf{V}$. The unit left and right singular vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are the three column vectors of $\mathbf{U}$ and $\mathbf{V}$ respectively. The singular values σ_i are the diagonal elements of Λ. By convention they are ordered such that $|\sigma_1| \geq |\sigma_2| \geq |\sigma_3| \geq 0$.

The left singular vectors $\mathbf{u}_i$ and right singular vectors $\mathbf{v}_i$ are mutually orthogonal, or $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}\mathbf{V}^T = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. An SVD therefore decomposes the data matrix $\mathbf{D}$ into its orthogonal parts $\sigma_i \mathbf{u}_i \mathbf{v}_i^T$. For 3C seismic data $\sigma_i \mathbf{u}_i$ represents the seismogram that is recorded by the 3C receiver in the direction of the polarisation vector $\mathbf{v}_i$. It holds that $\sigma_i \mathbf{u}_i = \mathbf{D}\mathbf{v}_i$. The total seismic energy in the polarisation direction $\mathbf{v}_i$ and over the time interval $[T_1, T_N]$ is given by σ_i^2. It is a property of SVD that the polarisation vectors $\mathbf{v}_i$ give the directions of maximal σ_1^2, intermediate σ_2^2, and minimal σ_3^2 seismic energy.

EVD is the decomposition of the data-correlation matrix $\mathbf{D}^T\mathbf{D}$ into its eigenvalues and eigenvectors, yielding

$$\mathbf{D}^T\mathbf{D} = \mathbf{V}\Lambda^2\mathbf{V}^T = \sum_{i=1}^{3} \mathbf{v}_i \lambda_i \mathbf{v}_i^T. \tag{2}$$

This definition follows immediately from (1). The squared singular values give the eigenvalues $\lambda_i = \sigma_i^2$. The right singular vectors $\mathbf{v}_i$ are the eigenvectors of $\mathbf{D}^T\mathbf{D}$ and the left singular vectors $\mathbf{u}_i$ are the eigenvectors of $\mathbf{D}\mathbf{D}^T$.

3.1.1 Influence of isotropic random noise

There are a number of important assumptions to consider when using SVD or EVD to estimate the true signal $\sigma_S\mathbf{u}_S$ and polarisation $\mathbf{v}_S$ from data $\mathbf{D}$. The signal must be linearly polarised, and the noise must be isotropic and uncorrelated with the signal. Isotropic noise is not polarised and its energy is equal in all directions. If we assume that $\mathbf{D} = \mathbf{S} + \mathbf{N}$, with $\mathbf{S} = \sigma_S\mathbf{u}_S{\mathbf{v}_S}^T$ the true 3C signal and $\mathbf{N}$ the 3C noise, then we can rewrite (2) and obtain

$$\mathbf{D}^T\mathbf{D} = \mathbf{S}^T\mathbf{S} + \mathbf{N}^T\mathbf{N} = \mathbf{v}_S\lambda_S{\mathbf{v}_S}^T + \lambda_N\mathbf{I}. \tag{3}$$

The noise correlation matrix $\mathbf{N}^T\mathbf{N}$ for isotropic random noise is simply the product of the identity matrix $\mathbf{I}$ with the noise energy in any direction λ_N. The first eigenvector $\mathbf{v}_1$ of the data-correlation matrix $\mathbf{D}^T\mathbf{D}$ defines the direction of maximum energy in the data. This is automatically the signal polarisation and $\mathbf{v}_1 = \mathbf{v}_S$. The total energy in this direction is $\lambda_1 = \sigma_1^2 = \lambda_S + \lambda_N$, with $\lambda_1 > \lambda_2 = \lambda_3 = \lambda_N$. Since $\mathbf{D} = \mathbf{S} + \mathbf{N}$ we find that the waveform $\sigma_1\mathbf{u}_1 = \mathbf{D}\mathbf{v}_1 = (\mathbf{S} + \mathbf{N})\mathbf{v}_1$ is composed of the signal $\sigma_S\mathbf{u}_S = \mathbf{S}\mathbf{v}_1$ and the part of the random background $\mathbf{N}\mathbf{v}_1$. The signal-to-noise ratio SNR of this signal estimate $\sigma_1\mathbf{u}_1$

is thus

$$SNR = \frac{\lambda_S}{\lambda_N} = \frac{\lambda_1 - \lambda_2}{\lambda_2}. \tag{4}$$

3.1.2 Influence of non-isotropic noise

The assumption of uncorrelated isotropic background noise is unrealistic for most seismic applications. If the noise is polarised but uncorrelated with the signal then $\mathbf{N}^T\mathbf{N} \neq \lambda_N \mathbf{I}$, but still $\mathbf{S}^T\mathbf{N} = \mathbf{0} = \mathbf{N}^T\mathbf{S}$. Expression (3) for the data-correlation matrix $\mathbf{D}^T\mathbf{D}$ becomes

$$\mathbf{D}^T\mathbf{D} = \mathbf{S}^T\mathbf{S} + \mathbf{N}^T\mathbf{N} = \mathbf{v}_S \lambda_S \mathbf{v}_S{}^T + \sum_{i=1}^{3} \mathbf{v}_{N_i} \lambda_{N_i} \mathbf{v}_{N_i}{}^T. \tag{5}$$

The noise correlation matrix $\mathbf{N}^T\mathbf{N}$ depends on the noise eigenvectors $\mathbf{v}_{N_i}$ and the noise eigenvalues λ_{N_i}, with $\lambda_{N_1} \geq \lambda_{N_2} \geq \lambda_{N_3} \geq 0$. In general, the first eigenvector $\mathbf{v}_1$ of the data-correlation matrix $\mathbf{D}^T\mathbf{D}$ and the signal polarisation vector $\mathbf{v}_S$ will be different. The angle between both vectors is typically a function of the noise and signal polarisations and of the ratios between signal and noise eigenvalues (Bataille and Chiu, 1991; Souriau and Veinante, 1975).

We illustrate the problems with polarised noise in Figure 1. For simplicity we use a 2C example, but the inferences made, are true for any number of components. Correlation matrices are square, symmetric and have positive eigenvalues. They can be represented graphically as ellipses (or ellipsoids for more dimensions), which are their geometric forms. Figure 1A (left) shows the geometric forms of the data-, signal- and noise-correlation matrices, $\mathbf{D}^T\mathbf{D}$, $\mathbf{S}^T\mathbf{S}$, and $\mathbf{N}^T\mathbf{N}$ for synthetic data with isotropic noise. The signal is linearly polarised and its correlation matrix is represented by a line. The geometric form of the isotropic noise-correlation matrix is a circle.

Figure 1A (right) displays the variation in signal-to-noise ratio SNR_θ and data variance, or data energy, λ_θ as a function of azimuth. For each azimuth θ we can define a vector $\mathbf{v}_\theta = [\cos(\theta), \sin(\theta)]^T$. The data energy is then $\lambda_\theta = |\mathbf{D}\mathbf{v}_\theta|^2$. The signal-to-noise ratio $SNR_\theta = \frac{|\mathbf{S}\mathbf{v}_\theta|^2}{|\mathbf{N}\mathbf{v}_\theta|^2}$ is the ratio of signal energy to noise energy. Figure 1A (right) shows that for data with isotropic noise the signal polarisation $\mathbf{v}_S$ coincides with the first eigenvector $\mathbf{v}_1$ and with the direction of maximal signal-to-noise ratio $\mathbf{v}_{SNR}$. This $\mathbf{v}_{SNR}$ is the $\mathbf{v}_\theta$ for a θ that maximises SNR_θ. Note that $max(SNR_\theta) = SNR$ in (4).

Figure 1B gives the geometric forms of the correlation matrices and the signal-to-noise ratio SNR_θ and data-variance λ_θ as a function of azimuth θ for the case where the noise is not isotropic. Contrary to the isotropic case we find that the first eigenvector of the data-correlation matrix $\mathbf{v}_1$, the signal polarisation $\mathbf{v}_S$ and the direction of maximal signal-to-noise ratio $\mathbf{v}_{SNR}$ are all different. Therefore, the first left and right singular vectors $\mathbf{u}_1$ and $\mathbf{v}_1$ provide poor estimates for the signal $\sigma_S\mathbf{u}_S$ and its polarisation $\mathbf{v}_S$. We also point out that the isotropic definition of the signal-to-noise ratio SNR in (4) no longer holds

3.2 Weighted decompositions

It is possible to correct for the presence of polarised random noise by applying a weighting to the data $\mathbf{D}$ (Samson, 1983a,b). We define the weighted data matrix $\mathbf{D}^w = \mathbf{D}\mathbf{W}^{-0.5} = (\mathbf{S} + \mathbf{N})\mathbf{W}^{-0.5}$ and its correlation matrix $\mathbf{D}^{wT}\mathbf{D}^w$ for a weighting matrix $\mathbf{W} = \mathbf{N}^T\mathbf{N}$. When SVD (1) and EVD (2) are applied to the weighted data and its correlation matrix we get

$$\mathbf{D}^w = \mathbf{U}^w\Lambda^w\mathbf{V}^{wT} = \sum_{i=1}^{3} \mathbf{u}_i^w\sigma_i^w\mathbf{v}_i^{wT}, \tag{6}$$

and

$$\mathbf{D}^{wT}\mathbf{D}^w = \sum_{i=1}^{3} \mathbf{v}_i^w\lambda_i^w\mathbf{v}_i^{wT}. \tag{7}$$

From this point we will use the superscript 'w' to identify vectors $\mathbf{v}_i^w$ and $\mathbf{u}_i^w$, and values σ_i^w and λ_i^w that result from SVD and EVD on weighted data $\mathbf{D}^w$. Similar to (3) and (5) we can now rewrite the weighted data-correlation matrix $\mathbf{D}^{w\,T}\mathbf{D}^{w}$ as

$$\mathbf{D}^{w\,T}\mathbf{D}^{w} = \mathbf{W}^{-0.5}\mathbf{S}^{T}\mathbf{S}\mathbf{W}^{-0.5} + \mathbf{I} = \mathbf{v}_{S}^{w}\lambda_{S}^{w}\mathbf{v}_{S}^{w\,T} + \mathbf{I}. \tag{8}$$

This expansion demonstrates how the weighting has effectively transformed the polarised noise into isotropic noise, since $\mathbf{I} = \mathbf{W}^{-0.5}\mathbf{N}^{T}\mathbf{N}\mathbf{W}^{-0.5}$. Weighting affects the signal differently and the weighted signal $\mathbf{S}^{w} = \mathbf{S}\mathbf{W}^{-0.5}$ will have a different polarisation and amplitude. This follows from $\mathbf{S}^{w} = \sigma_{S}\mathbf{u}_{S}\mathbf{v}_{S}{}^{T}\mathbf{W}^{-0.5} = \sigma_{S}^{w}\mathbf{u}_{S}^{w}\mathbf{v}_{S}^{w\,T}$. Here, the unit vector $\mathbf{v}_{S}^{w}$ is the weighted signal polarisation and $\sigma_{S}^{w} = \sigma_{S}\left|\mathbf{v}_{S}{}^{T}\mathbf{W}^{-0.5}\right|$ the weighted signal amplitude. The signal remains unchanged such that $\mathbf{u}_{S} = \mathbf{u}_{S}^{w}$. As the noise in the weighted data has now become isotropic, it will align with the direction of maximal weighted signal-to-noise ratio $\mathbf{v}_{SNR}^{w}$ and the first eigenvector $\mathbf{v}_{1}^{w}$ with the polarisation vector $\mathbf{v}_{S}^{w}$ of the weighted true signal (Figure 1C). The signal-to-noise ratio of the weighted signal estimate $\sigma_{1}^{w}\mathbf{u}_{1}^{w}$ is found similarly to (4) and yields

$$SNR^{w} = \frac{\lambda_{1}^{w} - \lambda_{2}^{w}}{\lambda_{2}^{w}}. \tag{9}$$

Equations (6) and the property that $\mathbf{D} = \mathbf{D}^{w}\mathbf{W}^{0.5}$ lead to a new decomposition for the data matrix

$$\mathbf{D} = \sum_{i=1}^{3}\mathbf{u}_{i}^{w}\sigma_{i}^{w}\mathbf{v}_{i}^{w\,T}\mathbf{W}^{0.5} = \sum_{i=1}^{3}\hat{\mathbf{u}}_{i}\hat{\sigma}_{i}\hat{\mathbf{v}}_{i}{}^{T}, \tag{10}$$

We use the unit vectors $\hat{\mathbf{u}}_{i}$ and $\hat{\mathbf{v}}_{i}$ to represent the waveforms and the polarisations of the signals in the data $\mathbf{D}$. Their amplitudes $\hat{\sigma}_{i}$ are equal to $\sigma_{i}^{w}\ |\mathbf{u}_{i}^{w}\mathbf{W}^{0.5}|$. The polarisation vectors $\hat{\mathbf{v}}_{i}$ are not mutually orthogonal and different from the vectors $\mathbf{v}_{i}$ in (1). The same is true for the amplitudes $\hat{\sigma}_{i} \neq \sigma_{i}$. The signal vectors $\hat{\mathbf{u}}_{i}$ are orthogonal, but nevertheless different from the vectors $\mathbf{u}_{i}$ in (1). For completeness, (10) also leads to a new decomposition for the data-correlation matrix

$$\mathbf{D}^T\mathbf{D} = \sum_{i=1}^{3} \mathbf{W}^{0.5}\mathbf{v}_i^w \lambda_i^w {\mathbf{v}_i^w}^T \mathbf{W}^{0.5} = \sum_{i=1}^{3} \hat{\mathbf{v}}_i \hat{\lambda}_i {\hat{\mathbf{v}}_i}^T. \tag{11}$$

For signal and polarisation estimation on data with non-isotropic noise these decompositions are superior to SVD (1) and EVD (2). Earlier in Figure 1C we discussed that the first weighted right eigenvector $\mathbf{v}_1^w$ is identical to the weighted signal polarisation $\mathbf{v}_S^w$. The accompanying weighted signal estimate $\sigma_1^w \mathbf{u}_1^w$ of the true weighted signal $\sigma_S^w \mathbf{u}_S^w$ has maximal signal to noise ratio. Because, $\hat{\sigma}_i \hat{\mathbf{u}}_i {\hat{\mathbf{v}}_i}^T = \sigma_1^w \mathbf{u}_1^w {\mathbf{v}_1^w}^T \mathbf{W}^{0.5}$ it follows that $\hat{\mathbf{v}}_1$ gives the true signal polarisation $\mathbf{v}_S$ and $\hat{\sigma}_1 \hat{\mathbf{u}}_1$ is the estimate for the true signal $\sigma_S \mathbf{u}_S$ with highest possible signal-to-noise ratio as per (9).

To summarise, noise weighted data decomposition provides us with the least noise biased estimate of the true signal and an accurate estimate of the true signal polarisation. The decomposition itself is achieved in three main steps:

- Estimate the weighting matrix $\mathbf{W}$. In practice this is done using pre-arrival seismic data, and by assuming that the noise polarisation properties are stationary over the length of the experiment. Previous efforts have shown that weighting is robust for small errors in the estimated noise-correlation matrix (Du et al., 2000; Samson, 1983b),
- Weight the data and determine the Singular Value Decomposition of the weighted data $\mathbf{D}^w$ (10),
- Apply the inverse weighting matrix to the SVD of $\mathbf{D}^w$. This leads to the new decomposition for the original data $\mathbf{D}$, which is given in (10).

3.3 Decompositions on multiple stations

SVD and EVD naturally apply to data with any number of components and (6) and (7), or (10) and (11) can be generalised to deal with multiple 3C recordings simultaneously (Bear and Pavlis, 1999). Each of these individual 3C data matrices can contain information from a single event that is recorded on multiple receivers, or from repeat source data on the same receiver, or a combination of both. A new N *by* $3k$ data matrix $\mathbf{D}$ can then be generated by combining the k individual 3C data matrices $\mathbf{D}^j$, with $j = 1, ..., k$, such that $\mathbf{D} = [\mathbf{D}^1, ..., \mathbf{D}^k]$. We assume that the signal has identical waveforms $\mathbf{u}_S^j$ in each of the included datasets $\mathbf{D}^j$, but amplitudes σ_S^j and polarisations $\mathbf{v}_S^j$ may vary. A moveout correction or time-alignment between data from each 3C dataset is usually required. The main difference between the single and multi-station approach is that the summation in equations (6), (7), (10) and (11) must now be done over $3k$ instead of 3 and that the polarisation vectors $\hat{\mathbf{v}}_i$ have $3k$ rather than 3 elements. The new $3k$ by $3k$ weighting matrix $\mathbf{W}$ is calculated from an N *by* $3k$ matrix with pre-arrival data from all 3C stations.

The ultimate aim is to interpret the multi-station estimates $\hat{\sigma}_i$, $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{v}}_i$, with $i = 1, 3 \cdots, k$, in terms of signals $\hat{\sigma}_i^j \hat{\mathbf{u}}_i$ and polarisations $\hat{\mathbf{v}}_i^j$ at the individual $j = 1, \cdots, k$ 3C stations. To do this we break up the $3k$-element polarisation vectors $\hat{\mathbf{v}}_i$ into k, 3-element vectors $\hat{\mathbf{v}}_i^j$ for each station j. The values $\hat{\sigma}_i$ are also redistributed over the k original stations such that $\hat{\sigma}_i^j \hat{\mathbf{v}}_i^j = \hat{\sigma}_i [\hat{v}(3j-2)_i, \hat{v}(3j-1)_i, \hat{v}(3j)_i]^T$, with $\hat{v}(3j)_i$ the $3j^{th}$ element of $\hat{\mathbf{v}}_i$. For each station this results in an alternative decomposition for the data matrix

$$\mathbf{D}^j = \hat{\mathbf{U}} \hat{\Lambda}^j \hat{\mathbf{V}}^{j^T} = \sum_{i=1}^{3k} \hat{\mathbf{u}}_i \hat{\sigma}_i^j \hat{\mathbf{v}}_i^{j^T}, \tag{12}$$

and for the data-correlation matrix

$$\mathbf{D}^{j^T} \mathbf{D}^j = \hat{\mathbf{V}}^j \hat{\Lambda}^{j^2} \hat{\mathbf{V}}^{j^T} = \sum_{i=1}^{3k} \hat{\mathbf{v}}_i^j \hat{\lambda}_i^j \hat{\mathbf{v}}_i^{j^T}. \tag{13}$$

Contrary to the EVD and SVD decompositions in (1) and (2), and the noise weighted decompositions in (10) and (11), where the data in each station is the sum of 3 signals, the multi-station decompositions above view the data as the sum of $3k$ different signals. The waveforms $\hat{\mathbf{u}}_i$ of these signals are the same for all stations, but their amplitudes $\hat{\sigma}_i^j$ and polarisations $\hat{\mathbf{v}}_i^j$ may vary.

The main advantage of this multi-station approach is that it benefits from a potential $\sqrt{j}$-fold increase in the signal-to-noise ratio, without the need to assume identical signal polarisations at all stations. The latter assumption is made when stacking matrices. Even the influence of signal-correlated noise will be reduced as long as this noise is not present in all stations, or out of phase. Obviously this leads to better estimates of the signal $\hat{\sigma}_S^j \hat{\mathbf{u}}_S \approx \hat{\sigma}_1^j \hat{\mathbf{u}}_1$ and their polarisations $\hat{\mathbf{v}}_S^j \approx \hat{\mathbf{v}}_1^j$.

Multi-station data decompositions are achieved by combining different 3C data-matrices of the same wave into one large data-matrix. It is important that arrival-time differences between the individual 3C datasets are corrected for. After decomposition with (1), or (10) of the combined matrix, we simply break up the multi-station polarisation vectors into 3C polarisation vectors that relate to the individual stations. This leads to a new decomposition (12) for each individual station and consisting of $3k$ uncorrelated signals $\hat{\sigma}_i^j \hat{\mathbf{u}}_i^j$ with non-orthogonal polarisations $\hat{\mathbf{v}}_i^j$, where k is the number of 3C stations in the analysis.

3.4 Analytic signals

The use of time-domain signals limits us to analyse linearly polarised signals only. Signals with elliptical polarisations can be analysed with the use of analytic signals (Vidale, 1986). Linear polarisation are treated as a subset of elliptical polarisations with infinite ellipticity. The use of analytic signals also allows us to use shorter time windows or even compute instantaneous polarisation attributes (Schimmel and Gallart, 2003). The single, multi-station, weighted, or unweighted equations (1), (2), (10), (11), (12) and (13) can be used as before, but with the transpose (T) replaced by the complex conjugate transpose (H). The analytic signals are formed by taking the original seismograms and adding their Hilbert Transform as the complex part. The new, complex data matrix $\mathbf{D}^{\varsigma}$ is then simply formed by replacing the real seismograms with the analytic signals. The same applies to the construction of the complex weighting matrix $\mathbf{W}^{\varsigma}$. Note that from this point onward we will use the symbol ς to distinguish between real matrices and vectors, and their complex counterparts.

If we take the complex data matrix $\mathbf{D}^{\varsigma}$ for the noise-weighted multi-station case then the waveform estimates $\hat{\mathbf{u}}^{\varsigma j}_{i}$ and polarisation vectors $\hat{\mathbf{v}}^{\varsigma j}_{i}$ will be complex, but the amplitudes $\hat{\sigma}^{\varsigma j}_{i}$ remain real. Each column in $\hat{\sigma}^{\varsigma j}_{i}\hat{\mathbf{u}}^{\varsigma j}_{i}\hat{\mathbf{v}}^{\varsigma j^{H}}_{i}$ is the product of the analytic signal $\hat{\sigma}^{\varsigma j}_{i}\hat{\mathbf{u}}^{\varsigma j}_{i}$ with an element of the complex polarisation vector $\hat{\mathbf{v}}^{\varsigma j^{H}}_{i}$. Typically these columns differ by a constant amplitude ratio and phase rotation. If we consider only the real part $\Re\left(\hat{\sigma}^{\varsigma j}_{i}\hat{\mathbf{u}}^{\varsigma j}_{i}\hat{\mathbf{v}}^{\varsigma j^{H}}_{i}\right)$, then its columns are still scaled and phase-rotated copies of one another. This gives rise to elliptical particle motion and leads to the decomposition of $\Re\left(\mathbf{D}^{\varsigma}\right)$ into its elliptically polarised parts $\Re\left(\hat{\sigma}^{\varsigma j}_{i}\hat{\mathbf{u}}^{\varsigma j}_{i}\hat{\mathbf{v}}^{\varsigma j^{H}}_{i}\right)$. The semi-major, $\mathbf{a}^{j}_{i}$, and semi-minor, $\mathbf{b}^{j}_{i}$, axes of the polarisation ellipse of each elliptical part result from a real SVD (1) on $\Re\left(\hat{\sigma}^{\varsigma j}_{i}\hat{\mathbf{u}}^{\varsigma j}_{i}\hat{\mathbf{v}}^{\varsigma j^{H}}_{i}\right)$. They can also be derived directly from the complex polarisation vector, $\hat{\mathbf{v}}^{\varsigma j}_{i}$, as explained in Samson and Olsen (1980) and Vidale (1986).

4 Automatic estimation of polarisation

The basis of our automated procedure for polarisation analysis is a weighted multi-station approach that uses the analytic signal. The automation deals mainly with the problem of selecting the optimal data-window $t \in [T_1, T_2]$ over which the signal polarisation is estimated. The aim is to produce good quality polarisations with meaningful uncertainty estimates. The optimal analysis window should exclude all pre-arrival and post-arrival noise, and samples contaminated by coherent noise with deviating polarisation properties, such as secondary arrivals. This can be achieved by iteratively removing samples from the analysis window that produce unreasonably large angle misfits relative to a measured provisional polarisation. Individual samples can be removed since the use of analytic signals allow us to estimate instantaneous polarisation attributes.

We define our sample misfit angles $\gamma(t)$ from the weighted data $\mathbf{D}^{\varsigma w}$ and its first left singular vector $\mathbf{v}^{\varsigma w}_1$:

$$\gamma(t) = \arccos \frac{|\mathbf{d}^{\varsigma w\,H}(t)\mathbf{v}^{\varsigma w}_1|}{|\mathbf{d}^{\varsigma w}(t)|}. \tag{14}$$

The $3k$-element data-vector $\mathbf{d}^{\varsigma w}(t)$ represents one time-sample of the complex data and is a row vector of $\mathbf{D}^{\varsigma w}$. The complex matrix $\mathbf{D}^{\varsigma w}$ contains the moveout corrected analytic signals from k 3C stations. The $\gamma(t)$-values can vary between 0° and 90°, and give the angles between each data-vector and the provisional unit polarisation vector $\mathbf{v}^{\varsigma w}_1$. The angle misfits are preferably derived from weighted data as weighting normalises the noise contribution of different stations and receiver components, and makes the misfit angles $\gamma(t)$ more sensitive to the presence of coherent noise.

The misfit angles are used to define the spherical variance υ^2, which is given by

$$\upsilon^2 = 1 - \sum_{t=T_1}^{T_2} w(t)\cos^2\gamma(t), \qquad \text{where} \qquad w(t) = \frac{|\mathbf{d}^{\varsigma w}(t)|^2}{\sum\limits_{t=T_1}^{T_2} |\mathbf{d}^{\varsigma w}(t)|^2}. \tag{15}$$

This definition varies slightly from Fisher et al. (2004) and Butler (1992) because the contribution from each sample is weighed according to its normalised energy $w(t)$, and because the data are complex. In practice, the spherical variance is computed from the eigenvalues $\lambda^{\varsigma w}_i$ (7) of the weighted data-correlation matrix $\mathbf{D}^{\varsigma w\,H}\mathbf{D}^{\varsigma w}$ by

$$\upsilon^2 = 1 - \frac{\lambda^{\varsigma w}_1}{\sum\limits_{i=1}^{all} \lambda^{\varsigma w}_i}. \tag{16}$$

For a confidence percentage α we define the interval $[-\mu_\alpha, +\mu_\alpha]$, in which α percent of the $\gamma(t)$-angles are expected to lie. After Fisher et al. (2004) we have

$$\mu_\alpha = \arcsin e_\alpha \upsilon \qquad \text{and} \qquad e_\alpha = \sqrt{-\ln(1-\alpha)}. \tag{17}$$

The optimal data window for polarisation analysis is found by iteratively removing poor quality samples from an initial data window. At each iteration, a provisionally array-based polarisation vector $\mathbf{v}^{\varsigma w}{}_1$ is obtained with misfit angles $\gamma(t)$ for the samples used. Poor quality samples are identified by misfit angles $\gamma(t)$ that lie outside the expectancy interval $[-\mu_\alpha, +\mu_\alpha]$. This process is repeated until no outliers are identified, or until a minimum number of remaining samples is reached. Like the initial data window, the confidence percentage α and the minimum number of samples in the optimised window are user-defined parameters. We find the confidence percentage α is generally over 85% for adequately sampled data. The typical minimum number of samples is equivalent to halve or one dominant wavelength.

The optimised data window is usually irregularly sampled in time and contains the most consistent samples within the initial data window. Small changes in this initial data-window are therefore unlikely to affect the final solution. Note that the statistics of our procedure use the $\gamma(t)$-values of an entire array. These array-based estimates are preferred over their station-based counterparts since they ‘average out’ possible non-Gaussian and bimodal behavior of the angle misfits from individual stations. This benefits accuracy and robustness in the presence of polarised background noise.

5 Practical implementation on P-wave array data

The automated technique is applied in a three-stage procedure to P-wave events from a 3C receiver array. The first stage obtains accurate relative P-wave arrival-times between the different 3C receivers. The second stage is our automated polarisation analysis. The final stage uses estimated linearities and uncertainties to perform quality control and to identify stations with poor results. Estimates from the identified bad stations are discarded and the complete three-stage process is repeated for the remaining stations.

To avoid introducing artificial phase effects when estimating the signal and polarisation, it is vital that the relative arrival-times between stations are known accurately. We use an iterative and coherency-based stack optimisation technique similar to that of Rowe et al. (2002). Manually picked arrival-times are used to generate a 3C stack called the pilot trace. Cross-correlation functions between the 3C pilot and the 3C data from each station provide P-wave arrival-time corrections at every station. A new 3C pilot trace is then created with updated arrival-times. This process is repeated until the arrival-time corrections converge to zero. Conversion usually occurs quickly (≈ 3 iterations). It is best to use noise weighted data $\mathbf{D}^w$ to minimise the negative effects of stations with high noise levels.

In the final stage, a complex polarisation estimate $\hat{\mathbf{v}}^j_1$ at each j^{th} 3C station is obtained from (12). We take as the P-wave polarisation, the semi-major axis $\mathbf{a}^j_1 = \left[a^j_1(1), a^j_1(2), a^j_1(3)\right]^T$ of the polarisation ellipse defined by the complex polarisation $\hat{\mathbf{v}}^j_1$. The P-wave polarisation azimuth θ_j and inclination ϕ_j in each station is given by:

$$\theta_j = \arctan\left(\frac{a^j_1(2)}{a^j_1(1)}\right), \tag{18}$$

$$\phi_j = \arctan\left(\frac{\sqrt{a^{j^2}_1(2) + a^{j^2}_1(1)}}{a^j_1(3)}\right). \tag{19}$$

We introduce two useful quality-control parameters for P-wave analysis. The first is P-wave linearity lin_j and the second a confidence interval$[-\epsilon_{\alpha_j}, +\epsilon_{\alpha_j}]$ on the P-wave polarisation, $\mathbf{a}^j_1$. After Claassen (2001) we define

$$lin_j = |\mathbf{a}^j_1|. \tag{20}$$

The linearity lin_j is 1 for perfectly linearly polarised P-waves and decreases to 0.5 for P-waves with circular polarisations. For P-waves with high linearities we can relate the spherical variance v^2 in (16) to an uncertainty estimate for the P-wave polarisation $\mathbf{a}^j_1$. If N is the number of samples in the optimised analysis window, then the angular confidence interval $[-\epsilon_{\alpha_j}, +\epsilon_{\alpha_j}]$ relates to the confidence percentage α.

Similar to Fisher et al. (2004) we have

$$\epsilon_{\alpha_j} = \arcsin \frac{e_\alpha v}{\sqrt{N}} \qquad \text{and where } e_\alpha \text{ is defined in (17).} \tag{21}$$

6 Examples

6.1 Synthetic P-wave example

The following tests on synthetic data compare our automated weighted multi-station polarisation analysis with unweighted and single station methods. They also verify the use and reliability of the proposed quality control parameters ϵ_{α_j} and lin_j. The synthetic data are from a 4-station 3C array with 1ms sampling interval (Figure 3). The P-wave is a 30Hz sine wave with exponential amplitude decay. P-wave energy and arrival-times are kept constant between stations. Stations 3 and 4 are both contaminated with vertically polarised coherent noise. At station 3 this noise is a delayed 25 Hz sine wave, and at station 4 it is a simultaneously arriving 30 Hz cosine wave. Both waves have exponential amplitude decays and cause nonlinear particle motion. Random background noise is added to all stations. On stations 2 and 4 the energy level of this noise on the vertical component is increased so it becomes polarised. All other relevant signal and noise parameters are given in Table 1. The noise is such that the signal-to-noise ratio decreases with station number.

In a first test we evaluate five alternative analysis methods by comparing polarisation estimation errors with respect to the known signal polarisations. The confidence intervals $[-\epsilon_{95_j}, +\epsilon_{95_j}]$ and linearities lin_j that accompany these polarisation estimates are also considered. Table 2 provides the average test results from 200 realisations with different added random background noise. Note that the synthetic data in this test are optimally time-aligned and therefore do not require a cross-correlation based data alignment. Next is a summary of the five polarisation analysis techniques used:

1. Single-station 3C approach with isotropic noise assumption as in equation (1) and similar to Vidale (1986).
2. Single-station 3C approach with noise weighting as in equation (10) and similar to Samson (1983b).
3. Multi-station approach with noise weighting as in equation (12).
4. Noise weighted single-station approach as in equation (10) and with window optimisation applied.
5. Noise weighted multi-station approach as in equation (12) and with window optimisation applied.

These techniques are tested on 150 ms long (initial) data-windows. In case of weighting, the weighting matrix is calculated from 100 ms of simulated pre-arrival noise. An example of this pre-arrival noise is shown on Figure 3. Analysis methods that apply window optimisation (methods 3 and 5), use an interval of $[-\mu_{90}, +\mu_{90}]$ to identify bad quality samples. The minimum number of samples in the optimised analysis window is 30.

The estimation errors reported in Table 2 indicate significant variations in accuracy between the different techniques and stations. The data in station 1 are composed of signal and isotropic random noise, and satisfy the requirements for 3C unweighted SVD. As expected, all tested methods accurately estimate the P-wave polarisations. The signal in station 2 is contaminated by polarised random background noise and satisfies the requirements for weighted analyses. The superiority of weighted approaches in this case is confirmed by comparing measurement errors from methods 1 and 2 in Table 2. The typical error for any of the weighted methods is 0° on average. It is 14.5° for the unweighted 3C estimate. The data in station 3 are composed of signal and later arriving correlated noise. Both, weighted and unweighted methods, cannot handle this situation if the entire 150 ms analysis window is used. The most accurate polarisation estimates are found with our window optimisation procedure (Methods 4 and 5 in Table 2). This successfully removes the secondary arrival from the analysis window. Station 4 records signal with polarised random background noise and simultaneously arriving coherent noise. None of the tested methods is expected to handle this situation perfectly. However, the errors on multi-station polarisations ($\approx 7^\circ$ for methods 3 and 5) are significantly smaller than those on single station estimates ($> 16^\circ$ for methods 1, 2 and 4).

Linearities lin_j from single-station analyses (methods 1, 2 and 4) are typically high ($\geq 90\%$), yielding a misleading confidence in the estimations. In fact, only linearities from our method (5 in list) correctly indicate the accuracy of the polarisation estimates. Similar observations can be made for the confidence intervals $[-\epsilon_{95_j}, +\epsilon_{95_j}]$ that accompany the polarisation estimates. These intervals are mainly a function of the signal-to-noise ratio at the station of interest and the number of samples used. The multi-station approach best separates the signal from background noise and produces more reliable uncertainties for the polarisations. Single-station polarisation analyses underestimate the amount of noise in the data and, as a result, the uncertainty intervals.

To summarise we find that all tested methods provide similar and accurate results on good quality data. On poor quality data with polarised noise we find that our technique is superior to any of the other tested methods. Our approach also provides the most reliable polarisation linearity lin_j and confidence ϵ_{95_j} estimates in any case.

The next synthetic test investigates the working of our method in the presence of relative arrival time errors between stations. This is to simulate the case where no cross-correlation data alignment is applied, or where the data does not allow a robust arrival time alignment. We use the same data as in the previous test, but impose random arrival time perturbations at each station. These perturbations have standard deviations of respectively 0s, 1s, 2s, 3s, 4s and 5s. A 5s arrival time error at individual stations translates in a 10s relative timing error between stations. This is significant in relation to the 33s (30Hz) dominant wavelength of the synthetic data.

Figure 4 shows how, for each station, the polarisation errors, confidence estimates ϵ_{95_j} and linearities lin_j respond to increasing arrival time errors. These average estimates and their error bars are calculated from 200 realisations with different added random background noise. In general, the figure indicates a gradual reliability deterioration of polarisations and their quality estimates with increasing arrival time errors. Mistimings are especially problematic in stations with correlated noise. This is evident from the changes in polarisation errors at stations 3 and 4. Contrary to the linearities lin_j, we find that the confidence estimates ϵ_{95_j} vary only slightly. Timing errors with standard deviations as low as 3 ms cause the linearities at stations 1 to 3 to drop from $98.5\%\pm1\%$ to $90\%\pm10\%$. At station 4 the linearities drop from $87.5\% \pm 2\%$ to $80\% \pm 20\%$. This test indicates that an accurate arrival time correction is vital to the proper working of our method. When such accurate time allignment is impossible the quality estimates, especially linearities, will nevertheless indicate that polarisation estimates are unreliable. Linearities provided by our polarisation analysis are therefore especially usefull to identify stations with poor data quality and/or poor arrival time allignment.

6.2 Real data example

6.2.1 Dataset

The real data used for testing our automated approach comes from a passive seismic monitoring experiment that was conducted in the Valhall oil field in 1998. This field is situated in the Norwegian sector of the Central North Sea Graben. The reservoir rock is a mechanically weak and highly porous chalk that is compacting in response to production. This compaction has lead to sea-floor subsidence of up to 4 m and is considered the main reason for microseismic activity within the reservoir overburden (Dyer et al., 1999). Between June 2^{nd} and July 26^{th} 1998 a CGG-SST500, six-level, 3C, 30 Hz geophone string was installed in well 2/8-A3B. The geophone stations were deployed at 20 m intervals between depths of 2000 m and 2100 m, just above the reservoir and are numbered from 1 to 6 with increasing depth (Figure 5). The data has a sampling rate of 1000 Hz. A total of 303 locatable events were recorded during the experiment

(Dyer and Jones, 1998). The locations of these events are calculated using a local velocity model, the P-wave and S-wave arrival-times, and the P-wave polarisations at each station along the vertical array. Because the Valhall experiment uses a linear vertical array, the P-wave polarisations serve as estimates for the source-receiver direction. They are absolutely vital to adequately constrain the source locations. In early work the P-wave polarisations were obtained manually using a single station, unweighted EVD on raw P-wave data (Dyer et al., 1999). We use these results as a benchmark.

The overall noise polarisation properties are investigated using pre-arrival data. For the top 4 stations we found that the noise is predominantly isotropically polarised. The bottom station, and station 5 to a lesser extent, experienced 5 to 10 times higher levels of mainly horizontally polarised noise. This noise is thought to be related to the unclamped weight, attached to the lowermost geophone station (*R. Jones, pers. comm.*). The signal and noise both have typical bandwidths between 5 Hz and 60 Hz and peak at 30 Hz. As a consequence, no improvement in signal-to-noise ratio by bandpass filtering the data was expected and no filter was therefore applied. A small number of events and stations were checked for the presence of coherent (signal-generated) noise in the P-wave coda. Such noise was found in a number of records. The P-wave and S-wave are separated sufficiently in time (0.5 s) such that the noise on the P-wave is mainly due to scatter and multipathing effects. The presence of both polarised random and coherent noise makes this dataset ideal to test our automated polarisation analysis.

6.2.2 Implementation and results

First, accurate relative P-wave arrival-times are calculated by repeated cross correlation with pilot traces (see Section 5). The manual arrival-times of Dyer and Jones (1998) served as a starting point.

For the automatic P-wave polarisation analysis the length of the initial analysis window was set to 60 ms, starting with the P-wave arrival-time. This length was chosen after visual inspection of the data and comprises approximately two signal periods. We judged that a minimal number of 25 samples for the optimised analysis window is sufficient to provide a representative polarisation estimate. An α value of 90% is used in the window optimisation to remove all samples with misfit values $\gamma(t)$ that lie outside the expectancy interval $[-\mu_\alpha, +\mu_\alpha]$. The weighting matrix **W** was calculated from 300 ms of pre-arrival data. All stations with linearities, lin_j, of less than 95%, or with a polarisation uncertainty estimate, $|\epsilon_{95_j}|$, larger than 6° were deemed unreliable and subsequently discarded.

A total of 1463 polarisation estimates were obtained using these parameters. Figure 6 displays the results for a representative microseismic event. It contains the original data for all six stations and the associated automatically determined P-wave polarisations. On hodograms, or particle motion plots, for stations 1, 2 and 5 there is clear evidence that the P-wave particle motion becomes nonlinear after approximately half a period. Station 4 has seemingly the best data quality, whereas station 6 suffers from increased background noise. The automatically interpreted polarisations are in good agreement with the general P-wave particle motion trends.

In Figure 7 we show the histograms of the polarisation uncertainty estimates $|\epsilon_{95_j}|$ for each station j. These polarisation uncertainty estimates have slightly asymmetric distributions and range between 0.5° and 6°. The shape and spread of these uncertainty distributions confirm that the threshold of 6° for identifying bad estimates is appropriate for this dataset. The median of the uncertainty estimates decreases from 3° in station 1 to 2.1° in station 4. Values then rise to 3.4° in stations 5 and 6. Station 1 is furthest away from the cluster of microseismic events and has lower overall signal-to-noise levels due to longer travel paths. This explains the progressively decreasing uncertainty estimates found in stations 1, 2, 3 and 4. This trend is not continued in results from stations 5 and 6 due to increased noise from the unclamped weight beneath station 6.

6.2.3 Comparison with manually determined polarisations

The accuracy and quality of our automatically obtained polarisation estimates is finally assessed by comparing them with the manual polarisations from Dyer and Jones (1998). We analyse the scatter of our polarisations with respect to the benchmark ones. The scatter estimate is independent of the range of polarisations and is evaluated per station and plotted in stereographic projection. First, we take the manually obtained reference polarisation for a given station and event, and rotate it through azimuth and inclination to the vertical, or pole (points *A* and *B* to *a* and *b* in Figure 8). The same rotation is applied to the relevant polarisation from our set of estimates (points *A*' and *B*' to *a*' and *b*' in Figure 8). All reference polarisations lie on the pole after rotation and our polarisations will scatter around it. An overall deviation of our rotated polarisations from the pole indicates a systematic difference between both sets of estimates. Scatter in the N-S direction, such as between *b* and *b*', relates to inclination differences between the two sets of polarisations. Scatter in the E-W direction, such as between *a* and *a*', relates to azimuth differences (Figure 8).

We determine the centre of mass of our rotated polarisation estimates and its 95% expectancy limits max_{95} and min_{95} (Fisher et al., 2004). If the differences between our and manually obtained reference estimates of Dyer and Jones (1998) are perfectly random then the angle between the centre of mass and the pole is zero and $max_{95} = min_{95}$. Ideally, the values for max_{95} and min_{95} should also be similar to the estimated confidence limits on our polarisations ϵ_{α_j}.

The set of measurements from Dyer and Jones (1998) contain 1325 determined P-wave polarisations, 1228 of which are also measured using our automated approach and can thus be compared. A stereographic projection for the scatter between both sets of polarisation estimates in each station is given in Figure 9. The number of compared estimates per station and the angle between the pole and the centre of mass of the rotated estimates with the 95% expectancy intervals are given in Table 3.

Except for station 5, we find trivial deviation angles between the pole and the scatter centre of gravity. For station 5 we have an average deviation of $3°$ to the East, The scatter is moreover elongated between $2°$W and $10°$E. This indicates that our 223 compared polarisation azimuths are shifted azimuthally clockwise with respect to the benchmark ones. The most plausible reason for this bias are increased horizontal noise levels due to the unclamped weight below station 6. The observed azimuthal shift in polarisations is then expected, since our polarisation analysis corrects for noise polarisation effects while the method for the benchmark polarisations does not. For the same reason we would anticipate a more pronounced effect at station 6. Though, this is not so due to the extremely poor data quality at times of increased noise. The 63 compared polarisation estimates at this station relate mainly to ‘quiet’ times. Therefore, we do not expect a systematic shift in azimuths as is the case.

95% of the compared automatically obtained polarisation estimates differ by less than $4°$ from the manual polarisations from Dyer and Jones (1998). This is supported by the expectancy intervals max_{95} and min_{95} for differences between both sets of polarisations (Table 3). At most stations we find small and similar values for max_{95} and min_{95} values. This indicates that the differences between both sets of polarisations are mainly due to random picking errors. Finally, we observe a good agreement between max_{95} and min_{95} values and estimated confidence intervals ϵ_{95_j} for the automatic polarisations. This gives us an increased confidence in the automatically derived error estimates.

7 Discussion

Our noise-weighted 3C array method for polarisation analysis and signal extraction uses analytic signals as input data to analyse signals with elliptical and linear polarisations (Vidale, 1986). The input matrices could also consist of time series, or their Fourier transforms or wavelet transforms (Flinn, 1965; Lilly and Park, 1995; Samson, 1973). We prefer the analytic signal as it is computationally more cost effective than other transforms. Additionally, it is stable over short analysis windows which allows the removal of individual samples from the data-window. Fourier transforms, and especially wavelet transforms are nevertheless better suited for broadband data where signal and noise separate in the frequency domain.

Noise weighting is included to deal with the bias introduced by polarised noise. It is an important feature for data with low signal-to-noise ratios and strongly polarised noise (Bataille and Chiu, 1991; Souriau and Veinante, 1975). Weighting relies on prior knowledge of the noise correlation matrix and the assumption of stationary noise properties. Typically, the weighting matrix is estimated from 5 to 10 signal periods of pre-arrival data (Du et al., 2000; Samson, 1983a). The method is robust for small errors in the weighting matrix. To verify if weighting improves the estimates the signal-to-noise ratio SNR^w (9) must be greater than its unweighted counterpart SNR (4).

To improve accuracy in the presence of signal-correlated noise we combine data from different 3C stations into one analysis (Bear et al., 1999). The method only works when all data relate to the same signal. The data must also be accurately corrected for arrival-time differences. When the noise is unrelated between k combined datasets, the signal and polarisation estimates will benefit from a $\sqrt{k}$-fold increase in signal-to-noise. The main advantage over stacking like Jurkevics (1988), or Earle (1999), is that this multi-station approach provides signal and polarisation estimates at each station and that the signal polarisations and amplitudes may vary from station to station.

Our automated procedure for P-wave polarisation analysis relies on a number of parameters. The first set of parameters control the optimisation of the analysis window. It includes a pre-arrival data window for noise weighting and the initial P-wave window. These are typically 5 to 10, and 2 signal periods long. We find that reasonable changes in window lengths have little influence on the final result. The α-value that defines the misfit expectancy interval $[-\mu_\alpha, +\mu_\alpha]$, and the minimum number of samples in the optimised analysis window are more influential. For $\alpha < 85\%$ the window optimisation nearly always converges to the predefined minimum number of samples. We obtained best results with α-values of approximately 90%

and a minimal sample population of approximately 0.5 to 1 signal period. The second set of parameters are the threshold values for linearity lin_j and uncertainty ϵ_{α_j}. They control the identification of events and stations with poor data quality. This ultimately leads to the removal of low quality data from the analysis. We feel that linearity lin_j should always exceed 95% and that the error estimate should be less than $10°$. When this is not the case the polarisation estimate should be regarded as unreliable.

8 Conclusions

We introduce a complex, weighted, 3C-array method to determine the waveform and polarisation of the recorded seismic signal at each 3C station. Our method is based on Singular Value Decomposition (SVD), or Eigen Value Decomposition (EVD) of the analytic signal and combines approaches from Vidale (1986), Samson (1983a) and Bear and Pavlis (1999). It can handle linear and elliptically polarised signals and minimises the influence of polarised noise on the estimated signal properties.

We have combined this method with an optimisation procedure for selecting the data analysis-window and adapt it for P-wave polarisation analysis. It automatically detects and removes stations with poor quality data for each event. It also provides a P-wave polarisation with error and linearity estimates. Tests on synthetic and real data show that the combination of linearity and a confidence interval of errors accurately indicate the quality of the P-wave polarisation estimate.

We implemented our fully automated method on the 1998 Valhall microseismic dataset and obtained 1463 P-wave polarisations estimates from 303 events and 6 receivers. These estimates typically have 95% confidence limits of $3°$. We statistically compared our results with those from a previous study by Dyer et al. (1999). The reference P-wave polarisations were obtained by applying a simple, single-station EVD over a manually selected data window. The difference between both sets of polarisations is mainly due to random picking errors. 95% of all polarisation differences lie within $4°$ and none exceed $15°$. This confirms that our automated method provides meaningful error estimates and polarisations that are of comparable quality to manual measurements. An additional advantage is naturally that the automated analysis takes only a few minutes for the entire dataset, compared to a time intensive manual interpretation.

A significant subset of events in receiver station 5 are contaminated with horizontally polarised noise which causes a shift of up to 10° in the azimuths from the reference polarisations in Dyer and Jones (1998). Our technique corrects for this bias and therefore leads to more accurate polarisation estimates. In turn this improves the quality of the estimated event locations.

9 Acknowledgements

The authors would like to thank ABB Offshore Systems, BP Norge, Shell Exploration and Production UK and Schlumberger Cambridge Research for funding this research through the Leeds Microseismicity Project. We are grateful to ABB and BP for providing the data and to Rob Jones for providing the initial suggestion for this work. We greatly appreciate one anonymous reviewer and Robert de Franco for their helpful comments and suggestions.

References

Bataille, K., Chiu, J., 1991. Polarization analysis of high-frequency, three-component seismic data. Bull. Seism. Soc. Am. 81 (2), 622–642.

Bear, L., Pavlis, G., 1999. Multi-channel estimation of time residuals from broadband seismic data using multi-wavelets. Bull. Seism. Soc. Am. 89 (3), 681–692.

Bear, L., Pavlis, G., Bokelmann, G., 1999. Multi-wavelet analysis of three-component seismic arrays: Application to measure effective anisotropy at Piñon Flats, California. Bull. Seism. Soc. Am. 89 (3), 693–705.

Butler, R., 1992. Paleamagnetism: Magnetic Domains to Geologic Terranes. Blackwell Scientific Publications, Ch. Statistics of Paleomagnetic Data.

Claassen, J., 2001. Robust Bearing Estimation for Three-component Stations. Pure and Applied Geophysics 158, 349–374.

de Franco, R., Musacchio, G., 2001. Polarization filter with singular value decomposition. Geophysics 66 (3), 932–938.

Du, Z., Foulger, G., Mao, W., 2000. Noise reduction for broad-band, three-component seismograms using data-adaptive polarization filters. Geophys. J. Int. 141, 820–828.

Dyer, B., Jones, R., 1998. 3D processing and interpretation of the microseismic data recorded during June and July 1998 in the Valhall Field. Tech. Rep. IR03/82, CSM Associates LTD.

Dyer, B., Jones, R., Cowles, J., Barkved, O., Folstad, P. G., 1999. Microseismic survey of a North Sea reservoir. World Oil 220 (3).

Earle, P., 1999. Polarization of the Earth’s teleseismic wavefield. Geophys. J. Int. 139, 1–8.

Fisher, N., Lewis, T., Embleton, B., 2004. Statistical analysis of spherical data. Cambridge university press, Ch. 6.3, pp. 160–178.

Flinn, E., 1965. Signal analysis using rectilinearity and direction of particle motion. Proceedings of the IEEE 53, 1874–1876.

Jurkevics, A., 1988. Polarization analysis of three-component array data. Bull. Seism. Soc. Am. 78 (5), 1725–1743.

Lilly, M., Park, J., 1995. Multiwavelet spectral and polarization analyses of seismic records. Geophys. J. Int. 122, 1001–1021.

Mao, W., Gubbins, D., 1995. Simultaneous determination of time delays and stacking weights in seismic array beamforming. Geophysics 60 (2), 491–502.

Maxwell, S., Young, R., Busso, R., Jupe, A., Dangerfield, J., 1998. Microseismic logging of the Ekofisk reservoir. SPE/ISRM 47276 .

Montalbetti, J., Kanasewich, E., 1970. Enhancement of teleseismic body phases with a polarisation filter. Geophys. J. R. Astr. Soc. 21, 119–129.

Reading, A., Mao, W., Gubbins, D., 2000. Polarization filtering for automatic picking of seismic data and improved converted wave detection. Geophys. J. Int. 147, 227–234.

Rowe, C., Aster, R., Philips, W., Jones, R., Borchers, B., Fehler, M., 2002. Using automated, high-precision repicking to improve delineation of microseismic structures at the Soultz geothermal reservoir. Pure appl. geophys. 159, 563–596.

Samson, J., 1973. Description of the polarization states of vector processes: Applications to ULF magnetic fields. Geophys. J. R. Astr. Soc. 34, 403–419.

Samson, J., 1983a. Pure states, polarized waves, and principal components in the spectra of multiple, geophysical time-series. Geophys. J. R. Astr. Soc. 72, 647–664.

Samson, J., 1983b. The spectral matrix, eigenvalues, and principal components in the analysis of multichannel geophysical data. Annals Geophysicae 1, 115–119.

Samson, J., Olsen, J., 1980. Some comments on the description of the polarization states of waves. Geophys. J. R. Astr. Soc. 61, 115–129.

Schimmel, M., Gallart, J., 2003. The use of instantaneous polarization attributes for seismic signal detection and image enhancement. Geophys. J. Int. 155, 653–668.

Schulte-Pelkum, V., Masters, G., Shearer, P., 2001. Upper mantle anisotropy from long-period P polarization. J. Geophys. Res. 106 (B10), 21917–21934.

Silver, P., Chan, W., 1991. Shear-wave splitting and subcontinental mantle deformation. J. Geophys. Res. 96, 16429–16454.

Souriau, M., Veinante, J., 1975. Three adaptive filters for the detection of body waves, with application to deep seismic soundings. Bull. Seism. Soc. Am. 65 (4), 995–1012.

Teanby, N., Kendall, J.-M., Jones, R., Barkved, O., 2004. Stress-induced temporal variations in seismic anisotropy observed in microseismic data. Geophys. J. Int. 156, 459–466.

Vidale, T., 1986. Complex polarization analysis of particle motion. Bull. Seism. Soc. Am. 76 (5), 1393–1405.

Table 1: Signal and noise parameters for the synthetic 3C array P-wave dataset in Figure 3. The signal-to-noise values (SNR) are not derived using equations 9, or 4, but obtained directly from the synthetic data specifications in this table.

Table 2: True error, linearity (lin_j) and confidence estimate (ϵ_{95_j}) for P-wave polarisation estimates, obtained using 5 alternative measuring techniques on synthetic data. The reported values are the averages of 200 realisations using the synthetic data described in Figure 3 and Table 1. Each confidence interval defines the opening angle of the confidence cone that has the estimated polarisation as axis. The estimates with window optimisation relate to a sample misfit acceptance interval of $[-\mu_{90}, +\mu_{90}]$ and a minimum number of 30 samples in the optismised window.

Table 3: Measured differences between the reference and our P-wave polarisation estimates at each station, based on a EVD of the rotated polarisations vectors shown in Figure 9. Large angles between the pole and the first eigenvectors indicate systematic differences between our polarisations and the reference polarisations. max_{95} and min_{95} give the minimum and maximum 95% expectancy intervals for the difference between both sets of polarisations.

Figure 1: Left: Geometric forms of the data-, signal- and noise-correlation matrices, $\mathbf{D}^T\mathbf{D}$ and $\mathbf{S}^T\mathbf{S}$ respectively, and $\mathbf{N}^T\mathbf{N}$ for a 2D example. Right: The signal-to-noise ratio SNR_θ and data-variance λ_θ as a function of azimuth θ. **A**: Isotropic random noise and linearly polarised signal. The major axis of the data-correlation ellipse, $\mathbf{v}_1$, is parallel to the signal polarisation vector $\mathbf{v}_S$ and the direction of maximal signal-to-noise ratio $\mathbf{v}_{SNR}$. **B**: Non-isotropic random noise and linearly polarised signal without weighting. The major axis of the data-correlation ellipse $\mathbf{v}_1$ is different from the signal polarisation vector $\mathbf{v}_S$ and the direction of maximal signal-to-noise ratio $\mathbf{v}_{SNR}$. **C**: Example **B** after weighting. The major axis of the weighted data-correlation ellipse $\mathbf{v}_1^w$ is parallel to the signal polarisation vector $\mathbf{v}_S^w$. The vectors $\mathbf{v}_{SNR}$ and $\hat{\mathbf{v}}_1$ can be derived from $\mathbf{v}_1^w$ and $\mathbf{W} = \mathbf{N}^T\mathbf{N}$.

Figure 2: Flow-chart of the processing procedure for automated P-wave polarisation analysis on array data. A three-step procedure is followed, starting with the arrival-time alignment of data from different stations and followed by polarisations analysis and quality control.

Figure 3: Synthetic 3C P-wave dataset from a 4-station array. For each station and from left to right we show the data, its signal part and its coherent and random noise parts. The signal is a 30 Hz sine wave arriving at time=0 ms. The coherent noise in station 3 arrives 35 ms after the signal and is a vertically polarised 25 Hz sine wave. The coherent noise on station 4 is a 30 Hz cosine wave and arrives with the signal. Signal and coherent noise have exponential amplitude decay. Stations 2 and 4 are also contaminated with vertically polarised random background noise. More information on data properties can be found in Table 1.

Figure 4: Results of a test that investigates the performance of our technique in the presence of arrival time errors between stations. The horizontal axis indicates the standard deviation of the random arrival time perturbation at individual stations. The reported values for polarisation error, linearity and confidence

intervals are the averages of 200 realisations using the synthetic data described in Figure 3 and Table 1. The error bars relate to halve the standard deviation of the average values.

Figure 5: (Left) Valhall location map with 3D inset reporting the locations of the micro-earthquake sources recorded during a 1998 experiment and after Dyer et al. (1999). (Right) general geometry overview of the 1998 experiment. The geophone stations were deployed at 20 m intervals between depths of 2000 m and 2100 m, and are numbered from 1 to 6 with increasing depth.

Figure 6: The left panels shows an example of a microseismic event recorded at Valhall. The receivers are numbered 1 to 6 from top to bottom. The P-wave and S-wave arrivals are picked on the 3C data from each station. The horizontal bold lines that start at the P-wave arrivals mark the initial time windows used for polarisation analysis. The left panel shows the horizontal and vertical hodograms, or particle motion plots, for each station. They represent the 60 ms of data as indicated by the bold markers on the left panels. These hodograms are overlain with the automatically measured P-wave polarisations.

Figure 7: Histograms per station for the polarisation uncertainty estimates, ϵ_{95_j}. The polarisation uncertainties gradually decrease from stations 1 to 4. They are largest for stations 5 and 6, probably due to the presence of an unclamped weight attached to station 6.

Figure 8: Method to compare our polarisations with the reference polarisations in stereographic projection. *A* and *B* (in black) are two reference polarisations. *A*' and *B*' (in dark and light gray) are our alternative estimates for *A* and *B*. We move *A* and *B* to the pole, creating points *a* and *b*, first by rotating them horizontally to an azimuth of $0°$ and then vertically to an inclination of $90°$. Next, the same rotations are applied respectively to our polarisation estimates *A*' and *B*', thereby creating the points *a*' and *b*' close to the pole. The scatter of points *a*' and *b*' around the pole defines the difference between our polarisations and the reference ones. Scatter in the E-W direction (*a* and *a*') relates to differences in azimuth. Scatter in the N-S

direction (*b* and *b'*) relates to differences in inclination.

Figure 9: Stereographic projections of the scatter between our polarisation estimates and the reference polarisation estimates from Dyer and Jones (1998). Each scatter plot is generated as explained in Figure 8. Except for station 5, no systematic difference occurs between the manual and reference polarisations as the scatter is typically less than 5°.

Synthetic data properties					
Station number		**1**	**2**	**3**	**4**
Signal:	Azimuth (deg)	55	60	65	65
	Inclination (deg)	30	35	40	45
	Energy	1	1	1	1
Coherent noise:	Azimuth (deg)	-	-	0	0
	Inclination (deg)	-	-	90	90
	Energy	0	0	0.5	0.66
	Time delay (ms)	0	0	35	0
Random noise:	Energy (North)	0.06	0.06	0.06	0.06
	Energy (East)	0.06	0.06	0.06	0.06
	Energy (vert)	0.06	0.5	0.06	0.5
Signal-to-noise energy ratio:		5.5	1.6	1.5	0.8

Table 1:

Station number		**1**	**2**	**3**	**4**
3C complex SVD (1):	Error (deg)	0	14.5	17	24.5
(no weighting)	lin_j (%)	99.9	99.8	99.5	96.1
	ϵ_{95_j} (deg)	2.5	4	3	2.5
3C complex SVD (2):	Error (deg)	0	0	17	16.5
(with weighting)	lin_j (%)	99.9	99.8	99.4	94.4
	ϵ_{95_j} (deg)	2.5	5	3	4
4x3C complex SVD (3):	Error (deg)	0	0	9.5	6.5
(with weighting)	lin_j (%)	99.3	99.2	96.6	85.8
	ϵ_{95_j} (deg)	3.5	5	4	4.5
3C complex SVD (4):	Error (deg)	0	0	2	16
(with weighting)	lin_j (%)	99.9	99.8	99.6	94.2
(optimised window)	ϵ_{95_j} (deg)	2.5	5.5	4	4
4x3C complex SVD (5):	Error (deg)	0	0	0	8.5
(with weighting)	lin_j (%)	98.4	98.2	98.4	87.6
(optimised window)	ϵ_{95_j} (deg)	3.5	6.5	4	6

Table 2:

95% expectancy intervals for differences between original and automatic polarisations						
Station number	**1**	**2**	**3**	**4**	**5**	**6**
max_{95} (deg)	4.6	4.6	5	3.7	5.6	3.2
min_{95} (deg)	3.5	3.6	3.8	3.1	2.2	2.7
Deviation from vertical of overall difference (deg)	1	0.5	1	1	3	0.5
Number of estimates	194	222	254	255	223	63

Table 3:

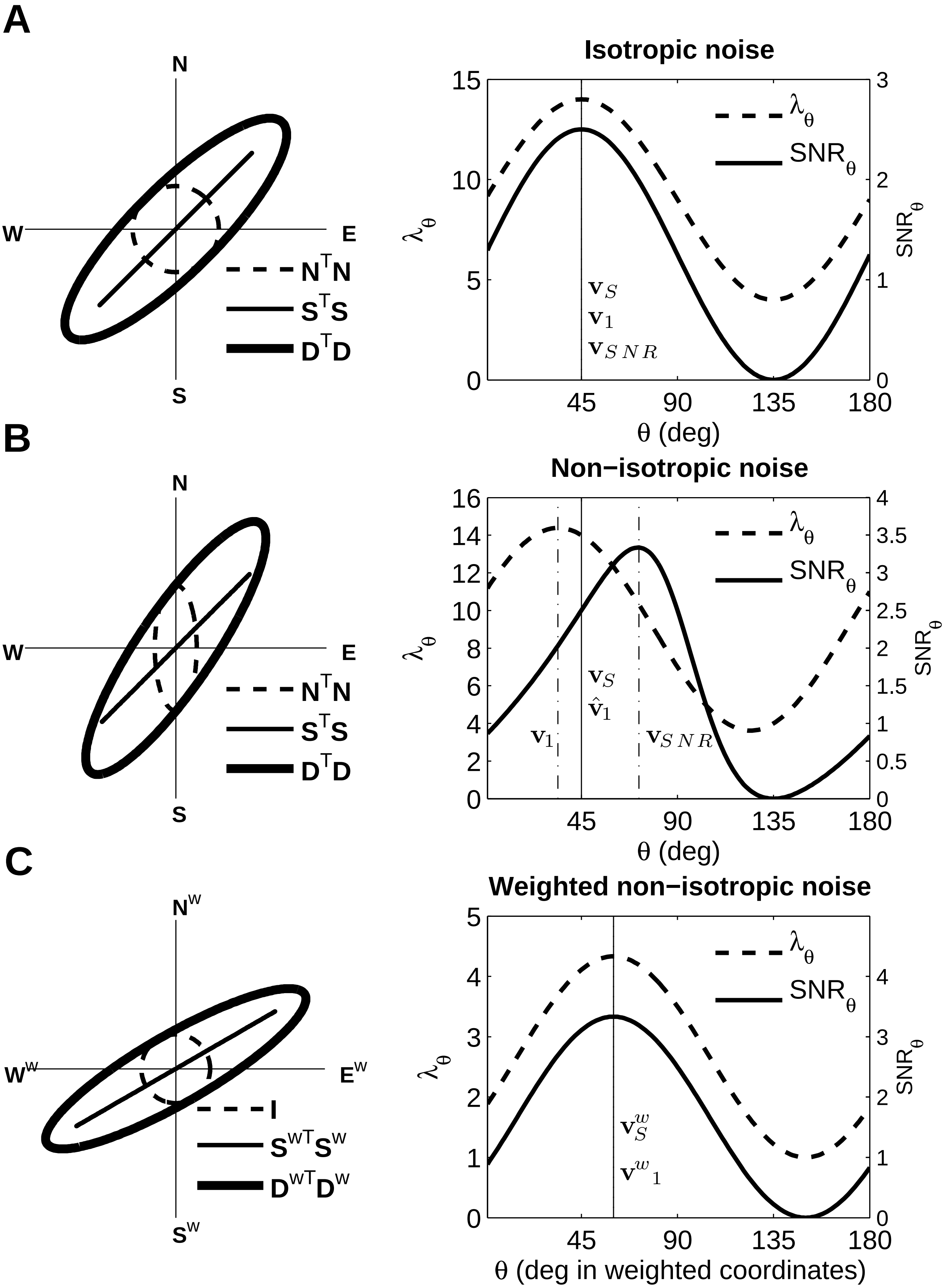

A
Isotropic noise
N
S
W
E
N^TN
S^TS
D^TD
λ_θ
SNR_θ
15
10
5
0
3
2
1
0
$\mathbf{v}_S$
$\mathbf{v}_1$
$\mathbf{v}_{SNR}$
45
90
135
180
θ (deg)
B
Non−isotropic noise
N
S
W
E
N^TN
S^TS
D^TD
λ_θ
SNR_θ
16
14
12
10
8
6
4
2
0
4
3.5
3
2.5
2
1.5
1
0.5
0
$\mathbf{v}_S$
$\hat{\mathbf{v}}_1$
$\mathbf{v}_1$
$\mathbf{v}_{SNR}$
45
90
135
180
θ (deg)
C
Weighted non−isotropic noise
N^w
S^w
W^w
E^w
I
$S^{wT}S^w$
$D^{wT}D^w$
λ_θ
SNR_θ
5
4
3
2
1
0
4
3
2
1
0
$\mathbf{v}_S^w$
$\mathbf{v}^w{}_1$
45
90
135
180
θ (deg in weighted coordinates)

Figure 1:

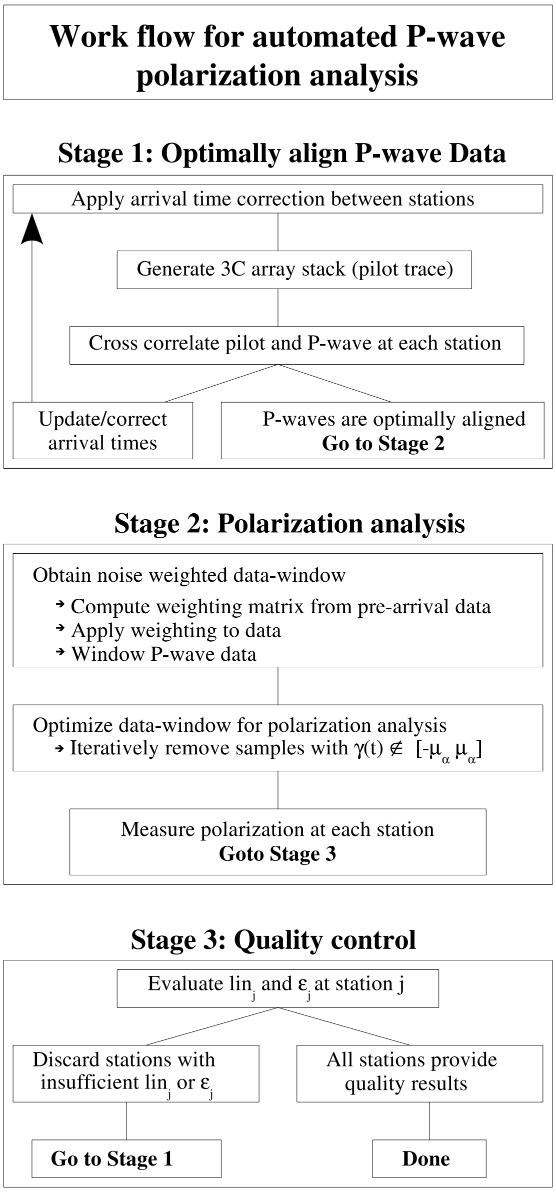

Work flow for automated P-wave polarization analysis
Stage 1: Optimally align P-wave Data
Apply arrival time correction between stations
Generate 3C array stack (pilot trace)
Cross correlate pilot and P-wave at each station
Update/correct arrival times
P-waves are optimally aligned
Go to Stage 2
Stage 2: Polarization analysis
Obtain noise weighted data-window
→ Compute weighting matrix from pre-arrival data
→ Apply weighting to data
→ Window P-wave data
Optimize data-window for polarization analysis
→ Iteratively remove samples with $\gamma(t) \notin [-\mu_\alpha \; \mu_\alpha]$
Measure polarization at each station
Goto Stage 3
Stage 3: Quality control
Evaluate lin_j and ε_j at station j
Discard stations with insufficient lin_j or ε_j
All stations provide quality results
Go to Stage 1
Done

Figure 2:

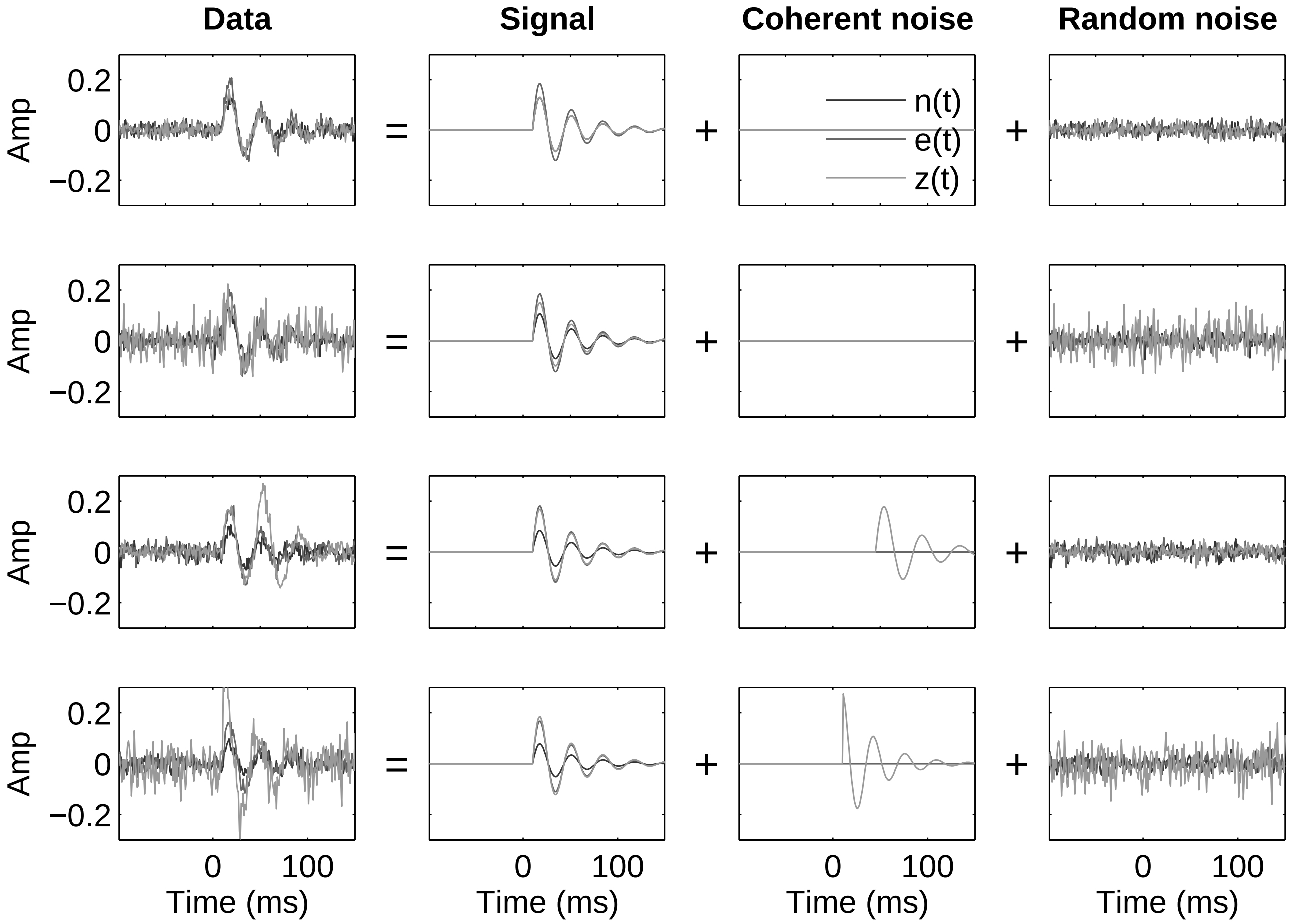

Data
Signal
Coherent noise
Random noise
0.2
0
−0.2
Amp
=
+
n(t)
e(t)
z(t)
0
100
Time (ms)

Figure 3:

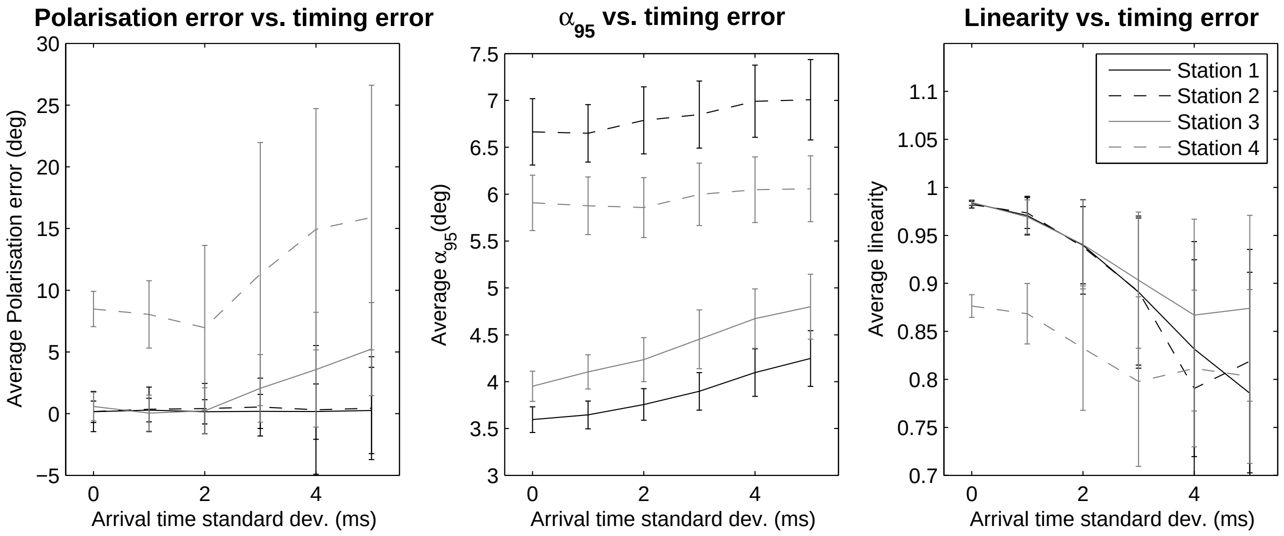

Polarisation error vs. timing error
Average Polarisation error (deg)
30
25
20
15
10
5
0
−5
0
2
4
Arrival time standard dev. (ms)
α_{95} vs. timing error
Average α_{95}(deg)
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3
0
2
4
Arrival time standard dev. (ms)
Linearity vs. timing error
Average linearity
1.1
1.05
1
0.95
0.9
0.85
0.8
0.75
0.7
0
2
4
Arrival time standard dev. (ms)
Station 1
Station 2
Station 3
Station 4

Figure 4:

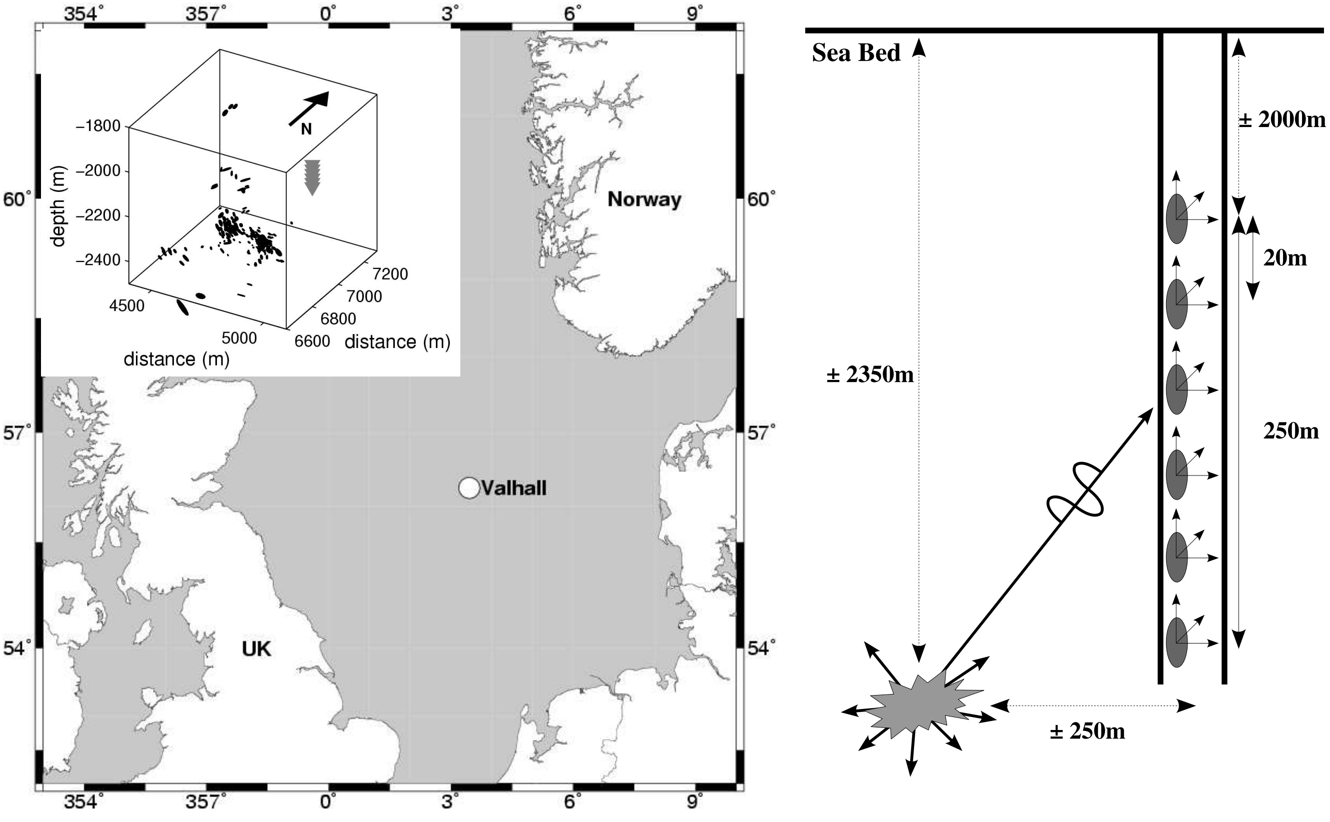
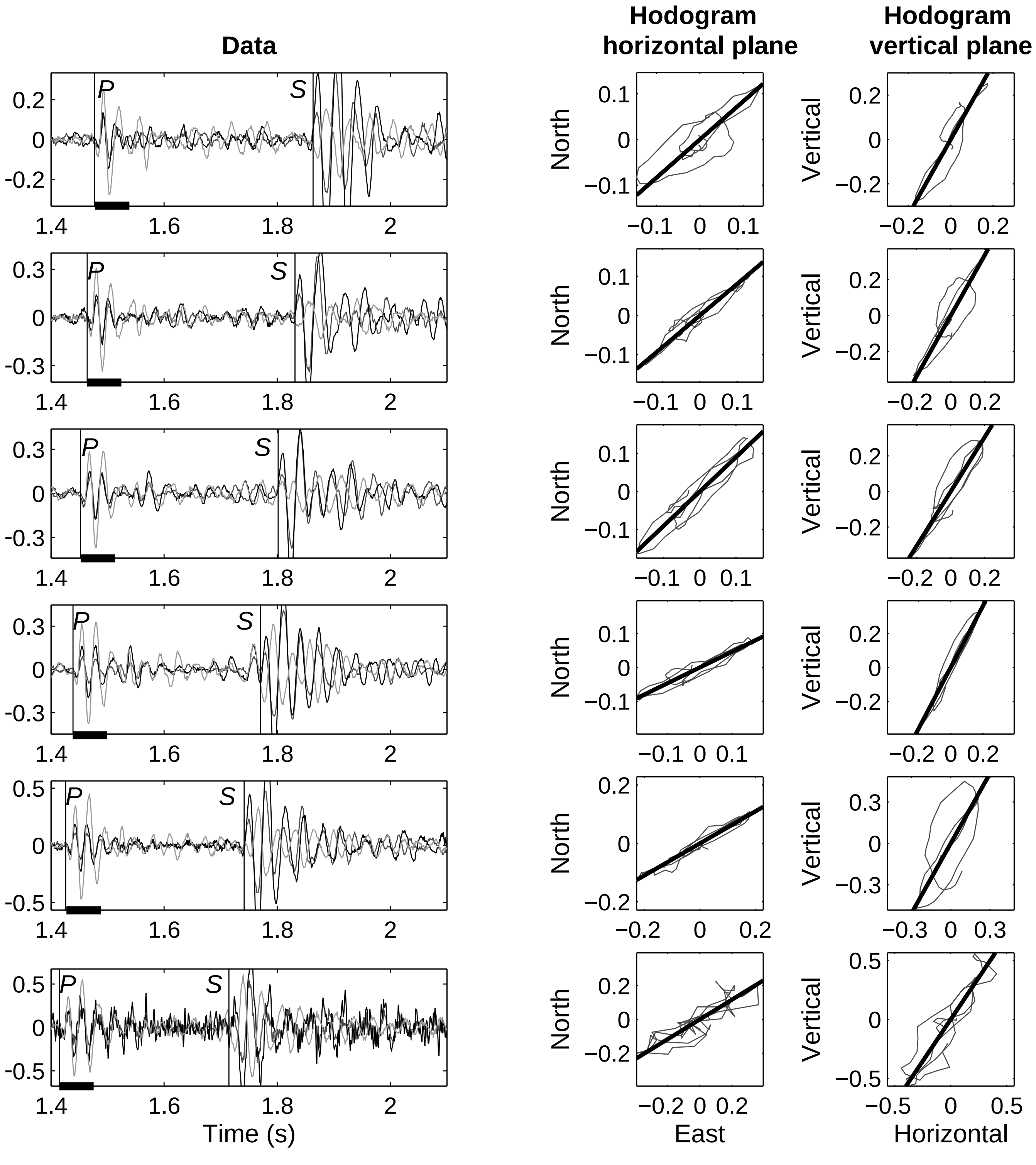

354°
357°
0°
3°
6°
9°
60°
57°
54°
−1800
−2000
−2200
−2400
depth (m)
N
4500
5000
distance (m)
6600
6800
7000
7200
distance (m)
Norway
Valhall
UK
Sea Bed
± 2000m
20m
± 2350m
250m
± 250m

Figure 5:



Data
Hodogram horizontal plane
Hodogram vertical plane
P
S
North
Vertical
East
Horizontal
Time (s)
1.4
1.6
1.8
2
0.2
0
-0.2
0.3
-0.3
0.5
-0.5
0.1
−0.1
−0.2
−0.3
−0.5

Figure 6:

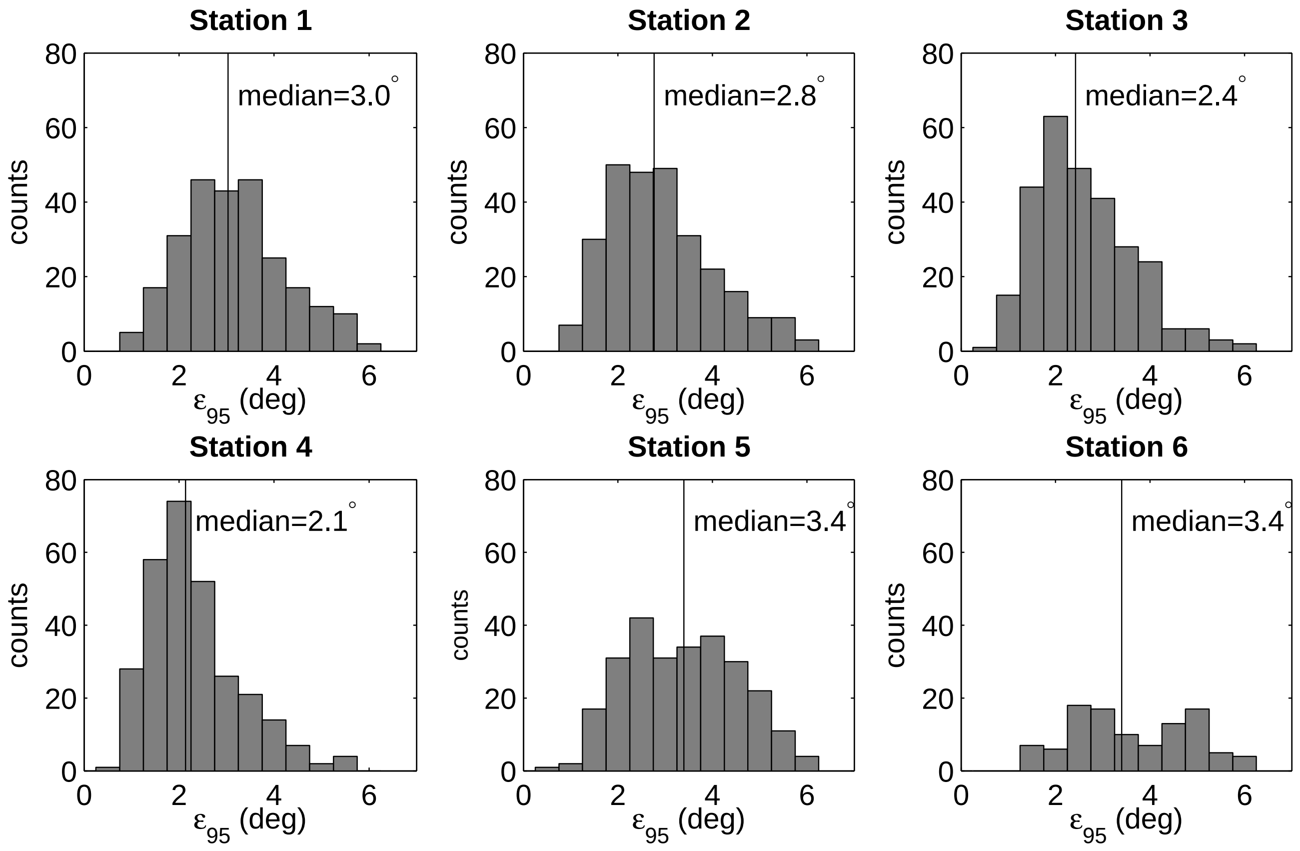

Station 1
Station 2
Station 3
Station 4
Station 5
Station 6
median=3.0°
median=2.8°
median=2.4°
median=2.1°
median=3.4°
median=3.4°
counts
ε95 (deg)
0
20
40
60
80
0
2
4
6

Figure 7:

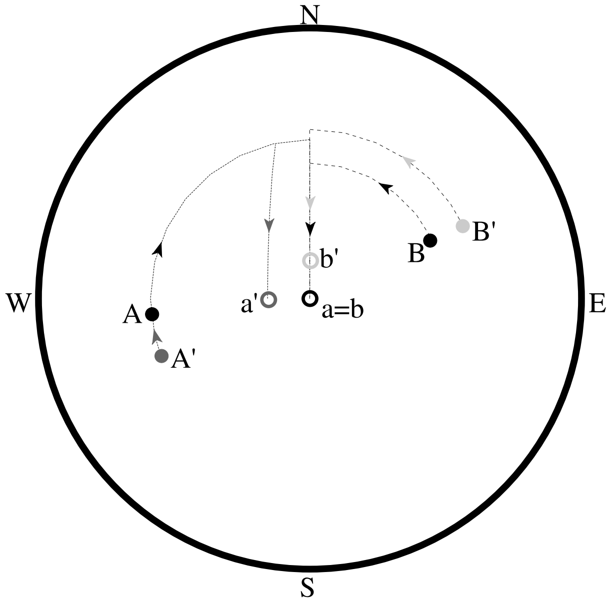


Figure 8:

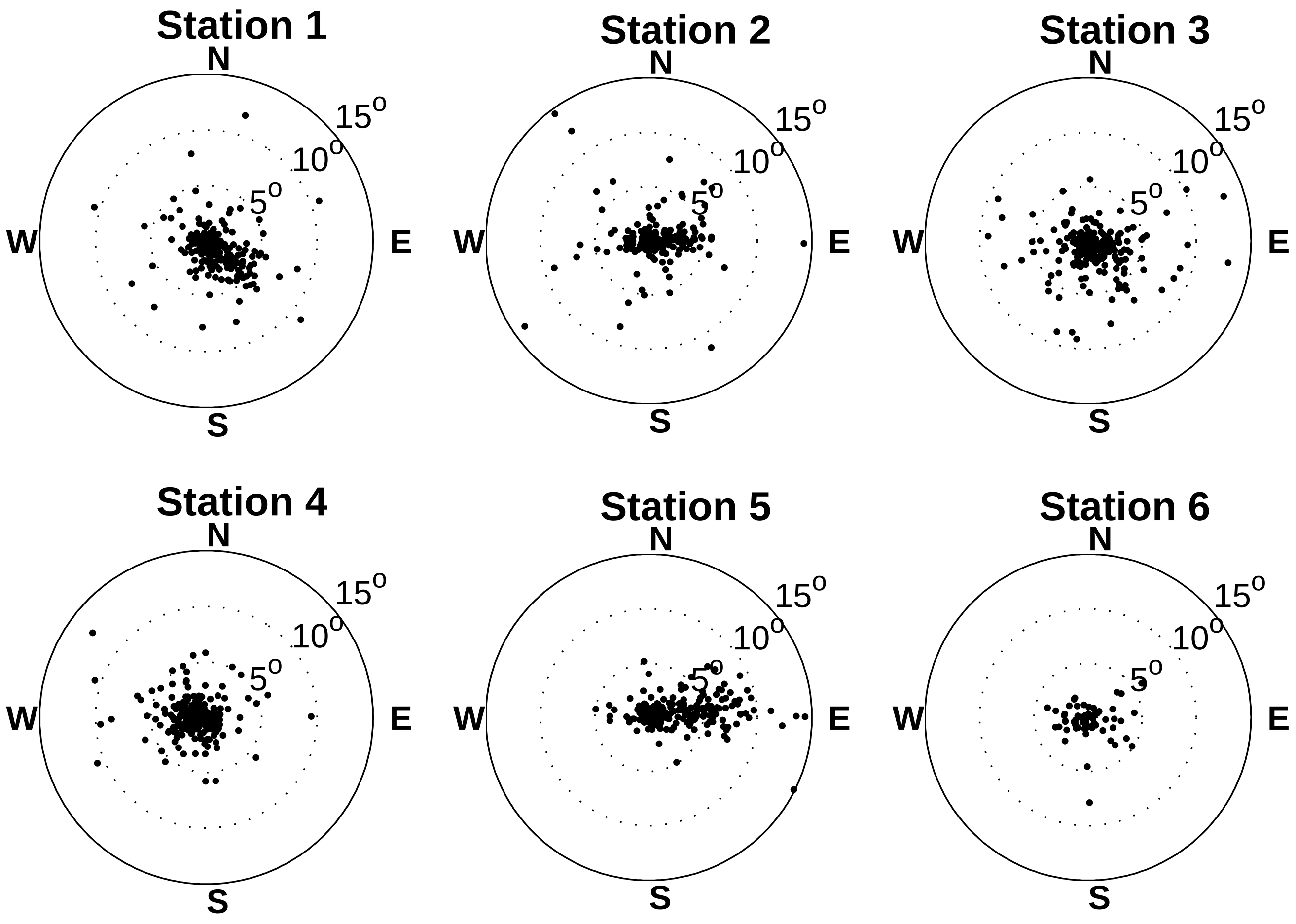

Station 1
Station 2
Station 3
Station 4
Station 5
Station 6
N
E
S
W
15°
10°
5°

Figure 9: