

A Dynamic Linear Model For Three-Component Seismic Waveforms

**Don Simone Daly, Kevin Anderson
Alan Rohay, Wes Nicholson
Pacific Northwest Laboratory**

Contract No. DE-AC06-76RLO 1830

Sponsored by the United States Department of Energy
Office of National Security and Nonproliferation

Technical Editor: Frank Ryan

Abstract

A three-component waveform is modeled as a sum of trigonometric functions using Dynamic Linear Modeling (DLM), a statistical extension of Kalman filtering. The DLM method converts a waveform into a multivariate time series of amplitude and phase-angle estimates. This time series is then transformed into a multivariate time series of descriptive features such as magnitude, direction-of-travel and dimensionality which are more closely related to the underlying seismic phases. There are many possible transformations, and several are considered here. This paper also explores the expected behavior of the dimensionality, magnitude and direction-of-travel time series over the duration of a phase. Dynamic Linear Modeling and parameter transformation is illustrated with an artificial waveform having compressional and Rayleigh phases.

Key words: Dynamic Linear Model, Kalman filter, waveform characterization, eigenanalysis.

Objective

The object of this study is to apply a trigonometric-based dynamic linear model (Kalman filter model) to a three-component seismic waveform in order to derive a set of descriptive features useful for CTBT discrimination.

Preliminary Research Results

Discriminating between a natural or manmade source of seismic activity begins with the extraction of descriptive features from a seismic waveform. Our approach to this extraction is straightforward. First, for a selected set of frequencies, model a three-component waveform as a sum of trigonometric functions with time-varying amplitude and phase-angle parameters. Then, transform the amplitude and phase-angle time series into a multivariate time series of features (e.g. dimensionality, direction-of-travel, magnitude) that are more descriptive of the underlying seismic phases. Next, analyze these descriptive series to identify and characterize the seismic phases and, to distill from them a set of statistics that summarize the discriminating information for a three-component waveform.

Three-component seismic waveforms have been described using a trigonometric-based Dynamic Linear Model (DLM), a statistical extension of the Kalman filter which is well-suited to the transient, nonstationary nature of seismic time series. In the DLM state space model, the estimates of time-varying amplitude and phase-angle parameters are strongly influenced by a three-component waveform's evolving cross-component and temporal covariance structure. Fitting the Dynamic Linear Model produces time series of parameter estimates and both cross-component and temporal covariance matrices.

One set of transformations of parameter estimates to time series of more relevant features involves the eigenanalysis of frequency-specific cross-component covariance matrices at each time step. More descriptive features (e.g. direction-of-travel, dimensionality and magnitude) are calculated from the eigenvalues and eigenvectors of each covariance matrix generating a multivariate time series of descriptive features.

The stability, or lack thereof, of the multivariate time series of descriptive features is directly related to the character of the underlying seismic phases. For example, the arrival of a P phase results in a sharp increase in eigenvalue magnitudes for the higher-frequency covariance matrices but minimal change in eigenvalue magnitudes for the lower-frequency covariance matrices. Related dimensionality and direction-of-travel time series remain relatively constant over the duration of the phase.

A Conceptual Model for a Seismic Waveform

Let us assume that a three-component seismic waveform results from the superposition of transient seismic phases and background noise. Let us further assume that the superpositioned waveform can be decomposed into a set of limited-frequency, three-component waveforms. For our purposes, a limited-frequency waveform has a waveshape whose frequency may vary over time in a band about a dominant frequency. Variation in frequency and amplitude of this waveshape is influenced by the arrival, direction-of-travel, dimensionality, polarity, magnitude and duration of certain underlying seismic phases; namely those seismic phases whose frequency content falls within the frequency band of the limited-frequency waveform. Therefore, the stability, or lack thereof, of the waveshape's frequency and amplitude are indicative of the underlying phase.

A Seismic State Space Model

Let $\Theta_\omega(t)$ represent the state, at time t , of one wave in the set of the f limited-frequency waves whose sum produces one component of a three-component waveform. For ω in the frequency set Ω , let the state of the limited-frequency wave be defined by

$$\Theta_\omega(t) = A_\omega(t) \{ \sin[\omega t + \phi_\omega(t)] + i \cos[\omega t + \phi_\omega(t)] \}$$

The state of the wave $\Theta_\omega(t)$ at time t is summarized by the two parameters $A_\omega(t)$ and $\phi_\omega(t)$. Note that the amplitude and phase-angle parameters are allowed to vary with time. Hence, a small set of limited-frequency waves can adapt to fit the superpositioned wave.

Derivations follow more easily if the state of the wave, $\Theta_\omega(t)$, is expressed in matrix format. Let $\alpha_\omega(t) = \text{Re}[\Theta_\omega(t)]$ and $\beta_\omega(t) = \text{Im}[\Theta_\omega(t)]$. Then,

$$\Theta_\omega(t) = \begin{bmatrix} \alpha_\omega(t) \\ \beta_\omega(t) \end{bmatrix}. \quad (1)$$

Note that the amplitude $A_\omega(t)$ and phase-angle $\phi_\omega(t)$ can be determined from the real and imaginary parts of $\Theta_\omega(t)$ by

$$\begin{aligned} A_\omega(t) &= \sqrt{\alpha_\omega^2(t) + \beta_\omega^2(t)}, \\ \phi_\omega(t) &= \tan^{-1}\left(\frac{\alpha_\omega(t)}{\beta_\omega(t)}\right) - \omega t. \end{aligned}$$

Through a series of applications of trigonometric identities, $\Theta_\omega(t)$ can be defined as a recursive linear function of $\Theta_\omega(t-1)$. Namely,

$$\Theta_\omega(t) = \mathbf{G}_\omega \Theta_\omega(t-1)$$

where

$$\mathbf{G}_\omega = \begin{bmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{bmatrix}.$$

Using the representation $\Theta_\omega(t)$ for each state, the set of states of the f limited-frequency waves can be represented by a $2f \times 1$ matrix

$$\Theta'(t) = [\Theta'_{\omega_1}(t) \cdots \Theta'_{\omega_f}(t)].$$

Letting \mathbf{G} be a $2f \times 2f$ block-diagonal matrix with one block equal to \mathbf{G}_ω for each ω in Ω , the recursive definition can be extended to include $\Theta(t)$

$$\Theta(t) = \mathbf{G}\Theta(t-1).$$

Finally, the model, $\Theta(t)$, describing one component of a three-component waveform, is extended to accommodate three components by adding one column for each component to $\Theta(t)$ so that equation 1 becomes

$$\Theta_\omega(t) = \begin{bmatrix} \alpha_{\omega,n}(t) & \alpha_{\omega,e}(t) & \alpha_{\omega,v}(t) \\ \beta_{\omega,n}(t) & \beta_{\omega,e}(t) & \beta_{\omega,v}(t) \end{bmatrix}$$

where the subscripts n , e and v refer to the north, east and vertical components, respectively. The results listed above for the one component model apply directly to the three-component model.

Incorporating Stochastic Elements

In practice, $\Theta(t)$ is a function of its previous states plus any additional effects introduced by new phase arrivals, environmental factors, and/or random perturbations. Further, only sums of the real parts of the set of states $\Theta(t)$, say $\mathbf{Y}(t)$, are observed and these are accompanied by some observational error. The model can accommodate these elements with two equations; a system equation to describe the evolution of $\Theta(t)$ and an observation equation to describe the measurement of $\mathbf{Y}(t)$. These are

$$\begin{aligned} \Theta(t) &= \mathbf{G}\Theta(t-1) + \Psi(t) \\ \mathbf{Y}(t) &= \mathbf{F}'\Theta(t) + \nu(t), \end{aligned}$$

where $\mathbf{F}' = [1 \ 0 \ \dots \ 1 \ 0]$ sums the real parts of $\Theta(t)$; $\nu(t) \sim N(0, v\Sigma(t))$ accounts for the observation error; and $\Psi(t) \sim MVN(\mathbf{0}, \Gamma(t), \Sigma(t))$ allows for environmental uncertainties.

By choosing Normal and MultiVariate Normal distributions to model the uncertainties, we are assuming

that the uncertainties are as likely to be positive as negative, and, more likely near zero than not. We also assume that the uncertainties from time to time are independent, and that observation errors are independent of environmental uncertainties.

Fitting the Stochastic Model

The set of amplitude and phase-angle parameters, and the two covariance matrices at each time t are estimated by Dynamic Linear Modeling (Kalman filtering); a Bayesian approach is taken wherein the estimates at time t depend on the estimates at time $t-1$ and, to a lesser degree, on the estimates at times $t-2, t-3, \dots$ [WH89, MS83]. The estimation procedure makes extensive use of the temporal and cross-component associations in a waveform; associations induced by the underlying seismic phases. Further, under the Bayesian formulation, the parameter set at each time t has a joint probability distribution. Therefore, it is possible to make statements about the precision of the parameter estimates and, hence, about the estimates of phase characteristics derived from the parameters.

Eigenanalysis and Time Series of Descriptive Statistics

One set of transformations of amplitude and phase-angle estimates into statistics more closely related to seismic phase characteristics begins with the definition of limited-frequency signal covariance matrices and their decomposition into eigenvectors and eigenvalues. The definition of relevant cross-covariance matrices followed by eigenanalysis is common to waveform characterizations proposed by several authors [BRAN94, Jur88, SH91, Tak91, Vid86]. Time series of dimensionality, polarity and magnitude measures are derived from the eigenvalues. Direction-of-travel (incidence angle and azimuth) time series are derived from the eigenvectors.

Let $\hat{\alpha}_\omega(t)$ denote the three-component estimate $\hat{Y}_\omega(t) = [\hat{\alpha}_{\omega,n}(t) \ \hat{\alpha}_{\omega,e}(t) \ \hat{\alpha}_{\omega,v}(t)]$. The vector $\hat{\alpha}_\omega(t)$ is an estimate of the real parts of $\Theta_\omega(t)$ and a row from the seismic state space matrix $\Theta(t)$. For time t and frequency $\omega \in \Omega$, define the cross-component covariance $\Sigma_\omega(t)$ as a weighted average of $\Sigma_\omega(t-1)$ and the outer product of $\hat{\alpha}_\omega(t)$,

$$\Sigma_\omega(t) = \frac{w\Sigma_\omega(t-1) + \hat{\alpha}_\omega(t)\hat{\alpha}'_\omega(t)}{w+1}$$

Frequency-specific measures of waveform dimensionality, spatial magnitude and direction-of-travel for each time t can be calculated from the eigen-values and eigenvectors of $\Sigma_\omega(t)$. For instance, waveform dimensionality can be estimated by comparing the magnitudes of the eigenvalues $\lambda_\omega(t)$. If one eigen-value, say $\lambda_{\omega,1}(t)$, is significantly larger than the other two over a period of time, then the limited-frequency wave Θ_ω is linear during that interval, and characteristic of a compressional wave. The size of the eigenvalues $\lambda_\omega(t)$ is related to the energy in $\Theta_\omega(t)$ and, hence, to the magnitude of the underlying seismic phases which contribute to the composite wave at that frequency. One measure of magnitude is the logarithm of a weighted geometric mean of the eigenvalues. This measure is an extension of the M_b and M_s measures.

Direction-of-travel can be estimated from direction cosines, the entries in an orthonormal eigenvector. For the appropriate frequency and eigenvector, the arc cosine of the entry associated with the vertical component is a measure of the incidence angle. The arc tangent of the ratio of the east to north elements provides an estimate of the azimuth.

Though the value of these statistics at any time t can be interesting, it is the behavior of the time series of these statistics that is most informative. For instance, a significant change in the magnitudes of the eigenvalues should occur with the arrival of a P phase. Over its duration, the related dimensionality measure should be stable indicating linearity. Further, the appropriate direction-of-travel time series should also be stable during this period.

An Example

Application of seismic eigenanalysis to the DLM cross-frequency, cross-components covariance time series was reported earlier [BRAN94]. The DLM was fit to artificial waveforms, Little Skull Mountain earthquake aftershock waveforms, and NTS explosion waveforms recorded at Goldstone, California. Paralleling the by-

frequency eigenanalysis described in the previous paragraph, the cross-frequency covariance time series were transformed into sets of derived time series corresponding to waveform dimensionality, magnitude, incidence angle, and azimuth. The effectiveness of the DLM approach was then illustrated using these derivative time series. A highlight of the report was the introduction of a novel ternary plot to visualize a waveform's evolving dimensionality. The distinct dimensionality of different phases (Pn+Pg, P-coda and S-coda) was apparent using the ternary plotting technique.

The cross-frequency analysis [BRAN94] could not accurately extract phase-specific characteristics. Our current emphasis on frequency-dependent analysis is in response to this shortcoming. We have defined frequency-specific covariance matrices and are now re-analyzing the aforementioned waveforms. We have improved the selection of model frequencies, initialization parameter values and model-evolution constants. Preliminary results show that characteristics of individual phases are identified more easily using this modified frequency-specific eigenanalysis.

Artificial waveforms featuring compressional, shear and Rayleigh wave characteristics in combination with background noise were generated to tune and test the DLM methodology. A composite waveform with compressional and Rayleigh phases is presented in Figure 1. This figure also shows the DLM decomposition of the composite waveform into five limited-frequency three-component waveforms.

The compressional phase of the combination waveform is a series of seven impulsive spikes. The spikes were inclined from vertical 20, 20, 20, 20, 20, 30 and 20 degrees, respectively; and had azimuths of 290, 280, 290, 300, 290, 290 and 290 degrees, respectively. The Rayleigh phase is a low-frequency surface wave having a vertical component 90 degrees out of phase with the horizontal components and an azimuth of polarization set at 45 degrees.

A subset of the feature time series derived from the DLM fit to the artificial composite waveform is shown in Figures 2, 3 and 4. These figures illustrate the variety of useful information about seismic phases that can be extracted using the DLM model. In particular, the stability of certain properties is revealed. The sharp change in the feature time series with the arrival of phases is also apparent.

The DLM successfully separates, by frequency, the compressional and Rayleigh phases in the composite waveform, Figure 1. The Rayleigh waveform is apparent in the 1 Hz DLM reconstruction, while the compressional waveform is apparent in the the 2, 4, and 8 Hz DLM reconstructions.

An eigenanalysis of the frequency-specific cross-component covariances indicates elliptical dimensionality for the lowest frequencies (two co-dominant eigenvalues) and linear dimensionality for the higher frequencies (one dominant eigenvalue) — which is consistent with two-dimensional Rayleigh and one-dimensional compressional waveforms, respectively, as shown in Figure 2. The relationship between phase arrival and the frequency-specific eigenvalues is illustrated by the sharp increases in the 2, 4, and 8 Hz dominant eigenvalues with the arrival of the compressional pulses.

Incidence angles and azimuths can be derived from the direction cosines that make up the frequency-specific eigenvectors. The angles and azimuths derived from the 4 Hz eigenvectors relate to the compressional wave, Figure 3. Note the stability of the incidence angle and azimuth measures, and how well they track the true angles for the compressional phase (Note that the azimuth angles, 280, 290, and 300 degrees, are equivalent to -80, -70 and -60 degrees, as shown in Figure 3).

For a planar wave such as a Rayleigh wave, the plane of motion is defined by the eigenvectors associated with the two largest eigenvalues. These fluctuate over time. The normal to this plane is defined by the eigenvector associated with the smallest eigenvalue. This eigenvector is stable over time. Figure 4 presents the incidence angle and azimuth time series corresponding to the smallest 1 Hz eigenvalue and, hence, the normal to the Rayleigh plane of motion. These graphs show the most stable estimates related to the incidence angle and azimuth of the Rayleigh wave. The estimate of incidence angle of the normal is approximately 90 degrees indicating that the incidence angle of the Rayleigh plane is nearly vertical. The estimated azimuth of the normal is approximately 135 degrees indicating that the azimuth of the Rayleigh wave is 45 or 225 (135 +/- 90) degrees. Both these estimates agree with the parameters of the artificial Rayleigh phase.

Conclusions and Future Plans

For simple artificial waveforms, trigonometric-based Dynamic Linear Models effectively capture the spectral and dimensional characteristics of three-component waveforms in time series of amplitude and phase-angle

estimates. These time series are readily transformed into time series of descriptive statistics closely related to the underlying seismic phases. Graphs of these descriptive series display distinguishing features related to the type of underlying phase. For instance, over the duration of a P phase, the dimensionality time series for the appropriate frequency is stable, and indicates a linear form while the related time series of azimuths is nearly constant. Several of the descriptive time series contain sharp change points indicating a change in underlying phase (Figure 2). The DLM approach appears well suited to characterizing properties of seismic phases that are stable.

The next phase of our research will build upon the knowledge we have gained from the development and analysis efforts to date. We plan to address DLM adaptations, improvements, and developments in several areas:

- develop DLM modeling diagnostics to improve the DLM fit.
- develop user guidelines to aid implementation and interpretation of the DLM.
- develop guidelines for parameter initialization and frequency selection.
- extend classical methods to use the extensive information in the time series of coefficient and covariance estimates.
- derive statistical measures of precision/quality to accompany DLM-based phase characteristic estimates.
- improve DLM visualization tools.

References

- [BRAN94] D.K. Blough, A.C. Rohay, K.K. Anderson, and W.L. Nicholson. Dynamic Bayesian Filtering For Real-Time Seismic Analyses. In M.D. Denny, editor, *Symposium on the Non-Proliferation Experiment Results and Implications*, Livermore, CA, 1994. Lawrence Livermore National Laboratory. CONF-9404100.
- [Jur88] A. Jurkevics. Polarization Analysis of Three-Component Array Data. *Bulletin of the Seismological Society of America*, 78(5):1725-1743, 1988.
- [MS83] R.J. Meinhold and N.D. Singpurwalla. Understanding the Kalman Filter. *The American Statistician*, 37:123-127, 1983.
- [SH91] A. Suteau-Henson. Three-Component Analysis of Regional Phases at Noreas and Arcs: Polarization and Phase Identification. *Bulletin of the Seismological Society of America*, 81(6):2419-2440, 1991.
- [Tak91] T. Takanami. A Study of Detection and Extraction Methods for Microearthquake Waves by Autoregressive Methods. *Journal of the Faculty of Science, Hokkaido University*, 9(1):67-196, 1991.
- [Vid86] J.E. Vidale. Complex polarization analysis of particle motion. *Bulletin of the Seismological Society of America*, 76(5):1393-1405, 1986.
- [WH89] M. West and J. Harrison. *Bayesian Forecasting and Dynamic Models*. Statistics. Springer-Verlag New York Inc., New York, New York, 1989.

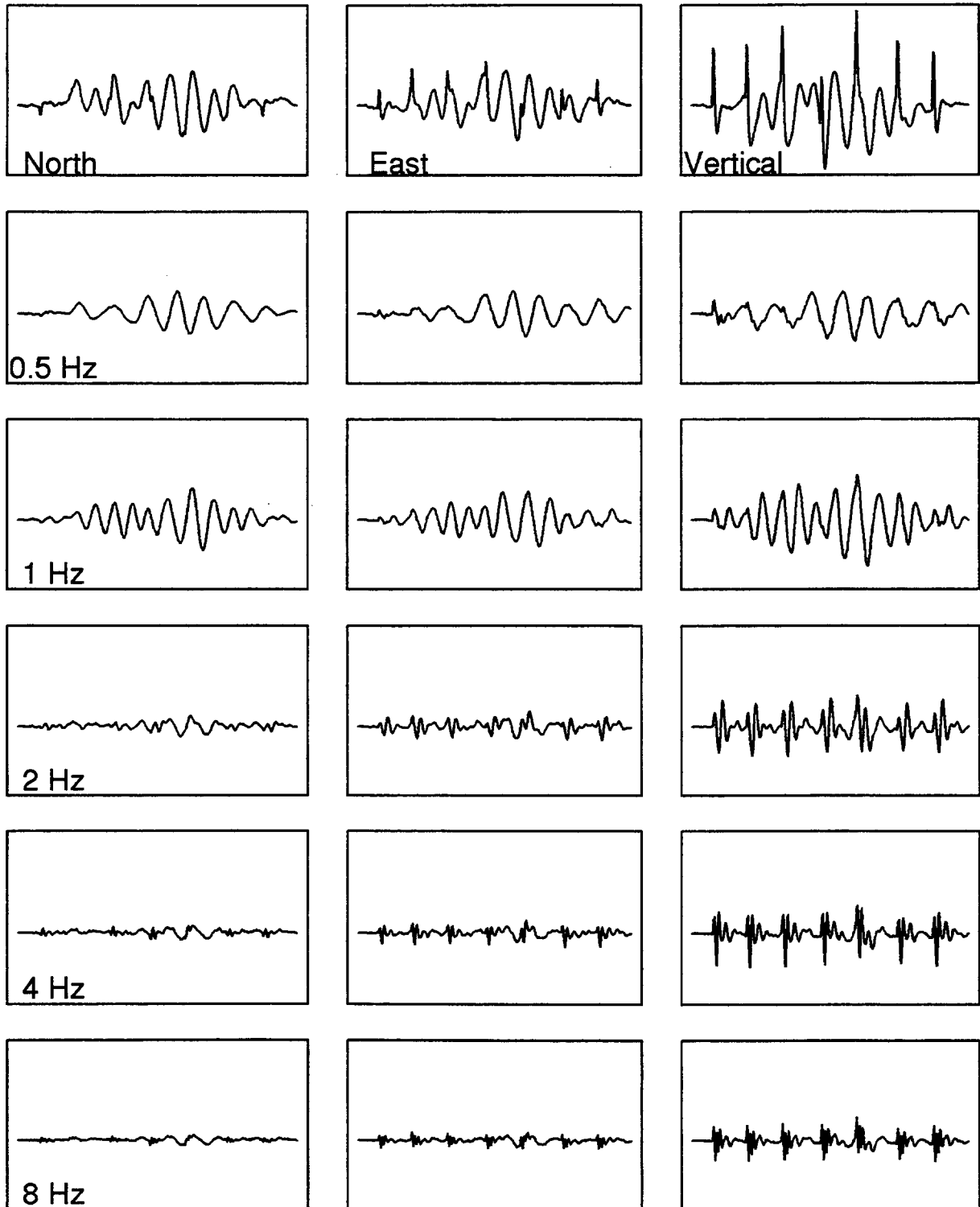


Figure 1: Artificial Three-component Waveform and DLM-reconstructed Components. Row 1 shows the north, east and vertical components of the artificial compressional and Rayleigh waveform. Rows 2 thru 6 show the separation of phases by frequency and waveform component achieved using a five-frequency DLM. The extracted Rayleigh phase is apparent in the 0.5 and 1 Hz panels, while the extracted compressional phase dominates the 2, 4, and 8 Hz panels.

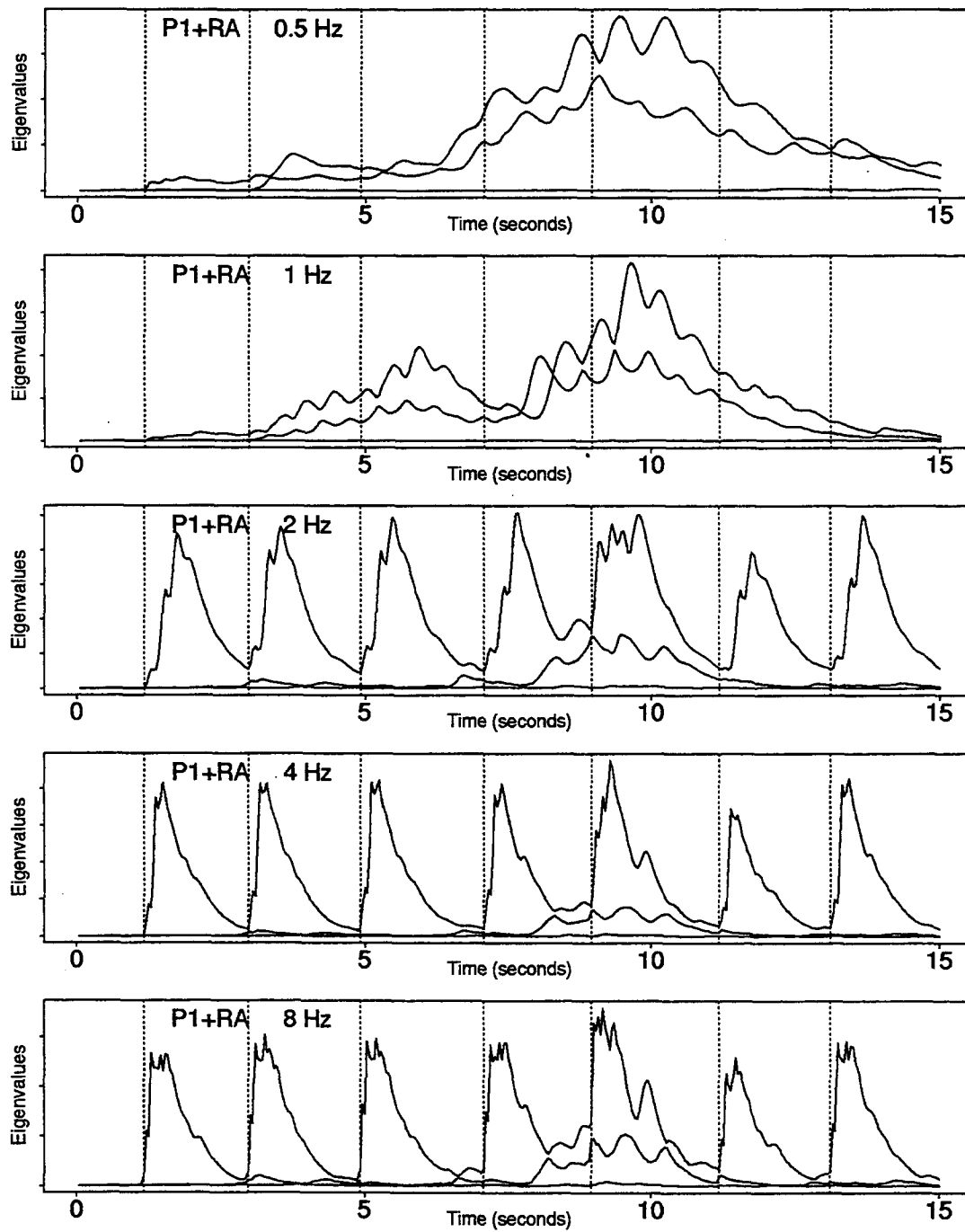


Figure 2: Time series of frequency-specific eigenvalues extracted from the DLM fit of the artificial composite waveform. At the lowest frequencies (top panels), the parallel profiles are consistent with the elliptical dimensionality expected with a Rayleigh wave. The bottom panels associated with the higher frequencies show the dominance of one eigenvalue, indicating linear dimensionality and a compressional wave. The sharp rise in the eigenvalue profiles in the lower panels is associated with the arrival of compressional waves. Note that the eigenvalue axis in each plot was scaled to aid visual comparison of within-plot profiles, so eigenvalue magnitudes are not comparable across plots.

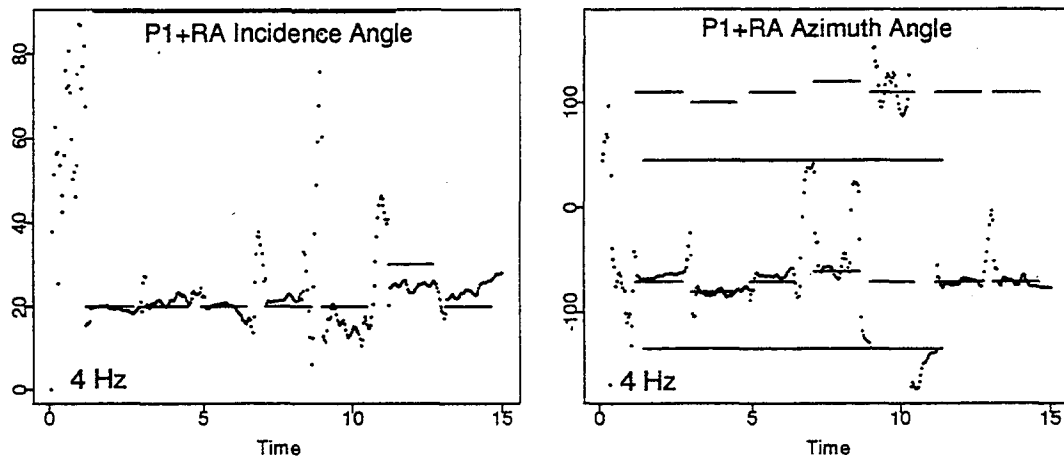


Figure 3: Time series of estimated incidence angles and azimuths for the “P1+RA” waveform, from the 4 Hz output of the DLM. The short and long horizontal bars denote approximately the known arrival times and durations of the compressional pulses and Rayleigh waveform, respectively. The rows of plots, top to bottom, are associated with the largest to smallest eigenvalues. At this frequency, the first row of graphs pertains to the one-dimensional compressional waveforms. The paired horizontal bars shown in the azimuth panels reflect the ambiguity in the eigenanalysis, which can resolve azimuths only to the angle plus or minus 180 degrees.

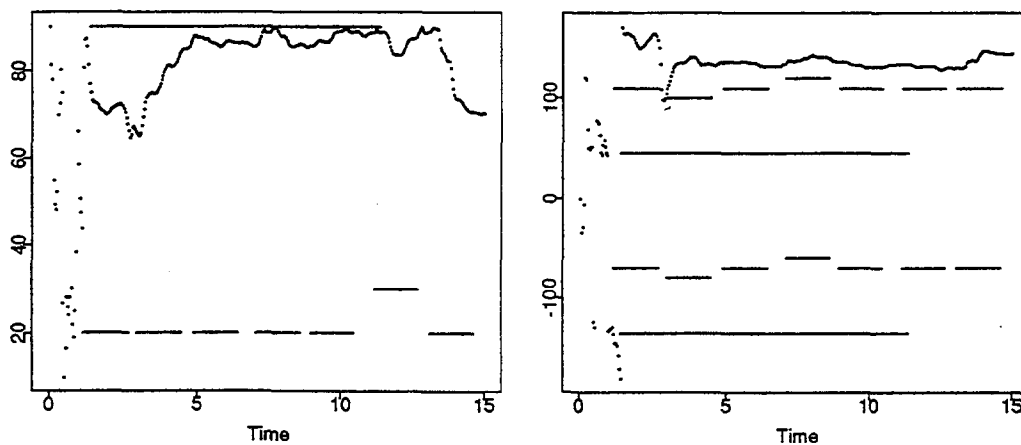


Figure 4: Time series of estimated “P1+RA” incidence angles (left) and azimuths (right) from the 1 Hz output of the DLM. The short and long horizontal bars denote approximately the known arrival times and durations of the compressional pulses and Rayleigh waveform, respectively. The rows of plots, top to bottom, are associated with the largest to smallest eigenvalues. At this frequency, the third row of graphs pertains to the dimension normal to the plane featuring the Rayleigh waveform. Note the stability of the features in these two panels. The incidence angle of the Rayleigh plane is nearly vertical, since this plane is orthogonal to the normal. Similarly, the azimuth is 45 degrees (90 degrees from the azimuth of the normal). The paired horizontal bars shown in the azimuth panels reflect the ambiguity in the eigenanalysis, which can resolve azimuths only to the angle plus or minus 180 degrees.