Autoregressive (AR) spectral estimates for Frequency-Wavenumber (F-k) analysis of strong-motion data

R. Rupakhety & R. Sigbjörnsson
Earthquake Engineering Research Center (EERC), University of Iceland

SUMMARY:
Frequency-wavenumber (F-k) spectra of seismic strong-motion array data are useful in estimating the backazimuth and apparent propagation velocity of seismic waves arriving at the array. Such estimates are required to model the wave passage effects while studying spatial variability of strong ground motions. Traditionally, F-k spectral computation is based on windowed periodogram estimates of spectral matrices. We present an alternative spectral estimate based on the well-known Autoregressive (AR) model. Such spectral estimates are found to be smoother than the windowed periodogram estimates, and can be directly used in F-k spectral analysis. We present an example application of the proposed technique using strong-motion data recorded by the SMART-1 array in Taiwan during the January 29 1981 M_L 6.3 earthquake. Our results, in terms of backazimuth and apparent propagation velocity, are found to be in excellent agreement with those reported in the literature.

Keywords: time series modelling, autoregressive model, seismic array, frequency-wavenumber spectra

1. INTRODUCTION

Earthquake ground motions recorded at different locations on the earth's surface vary in both amplitude and phase. Apart from overall attenuation of wave amplitudes with distance, various physical effects related to the seismic source, the wave propagation path, and the local site conditions result in spatially variable motion at the surface. Such spatial variability implies differential dynamic excitation at supports of structures spanning over large horizontal space, such as pipelines, bridges, dams, tunnels, transmission systems, etc. For these kinds of structures, the commonly invoked assumption of uniform support motion might lead to misleading interpretation of the response of actual structure to variable support motion. Example studies of such effects and their consequences can be found in Newmark and Rosenblueth (1971), Hindy and Novak (1980), etc. Zerva (2009) and the references therein contain a more detailed discussion on the engineering implications of spatially ground motions.

A simple and crude approach of introducing ground motion variability in structural analysis is to assume that the ground motion waves at supports are identical in amplitude and frequency content, but differ in phase due to their variable arrival times. This approach considers that the seismic wave travels through the horizontal dimensions of the structural base unchanged, and any difference in motion at different support points is entirely due to wave propagation. A more realistic modelling assumes variations in amplitude and/or frequency content, often referred to as incoherence of a random field. The quantification of the degree of incoherence is usually achieved by modelling of cross-spectral densities or their derivatives such as the so called lagged coherency (see Zerva 2009).

Spectral representation of a spatially variable ground motion field is, in practical applications, achieved by auto- and cross-spectral densities of ground motion recorded at closely spaced sensors of a ground motion array. Discretely recorded time series from the array are used in such spectral estimations. Fourier transform operations of time series lead to periodogram spectral estimates, which inherently involve large variance, which can be reduced, to a certain extent, by smoothing operations. Smoothing operations are, to a large degree, subjective in nature due to the fact that the selection of an optimum bandwidth of smoothing windows is based on experience and qualitative inferences rather than well-defined statistical criteria. This implies that different analysts working with the same set of data often obtain different spectral estimates (Broersen 2006).

Spectral estimators other than windowed periodograms exist in mathematical statistics. Over the last 30 years, parametric time series or model-based spectral analyses have gained popularity over traditional windowed periodogram approaches (see, for example, Kay and Marple 1981, Beamish and Priestley 1981). Parametric time series models (see Box et al 1994), e.g. autoregressive (AR), moving-average (MA), and autoregressive-moving-average (ARMA) models, are found useful in spectral analysis of experimental data collected in many fields of engineering, such as system identification (Ljung 1999), fluid mechanics (Pavageau et al. 2004), among many others. An application of AR models for short-period seismic noise appeared in Tjestheim (1975). Kozin (1988) discusses different aspects of fitting ARMA models to earthquake records. Leonard and Kennett (1999) used AR modelling of multi-component seismograms for spectral analysis and subsequent phase picking. Despite advances in the mathematical theory of AR and ARMA models and their widespread applications in several engineering applications, windowed periodogram still remain the primary tool in spectral analysis of strong-motion array data (see Zerva 2009).

Spectral estimates of seismic array data are useful, among other applications, in spatial variability modeling, study of source and wave propagation media, simulation of spatially variable ground motion. Wave propagation characteristics such as back azimuth and apparent propagation velocity of impinging waves at the site can be inferred from frequency-wavenumber (f-k) spectra conveniently derived from spectral estimates of seismic array data. The robustness and resolution of inferred quantities thus depend on the estimates of auto- and cross-spectral density estimates of seismic ground motion. Since spectral estimates obtained from AR (or ARMA) models provide better resolution than windowed periodogram estimates (see Leonard and Kennett 1999), the former are potentially more beneficial than the latter in f-k analysis of seismic array data. In this article we demonstrate the use of AR models in such applications. A brief overview of AR model calibration is presented along with introduction to f-k analysis method to estimate back azimuth and apparent propagation velocity. A case study using the data from the SMART 1 array in Taiwan in presented wherein the AR model is successfully applied to obtain an estimate of f-k spectra.

2. PARAMETRIC SPECTRAL ESTIMATES

Parametric spectral analysis is meant in this article to represent model-based time series analysis, where a suitable model is fitted to the seismic signal(s), and the parameters of the model are indentified by applying computational statistical techniques. Spectral estimates can then be expressed in terms of the parameters of the time-domain model. In time series analysis literature, parametric models refer to a class of time-domain models, such as AR, MA, and ARMA models, that are fitted to the signal. A comprehensive description of these different models is beyond the scope of this article, but can be found elsewhere (Kay and Marple 1981, for example). Spectral estimates obtained from such models are called as parametric spectral estimates in this work. In particular, the AR model is considered. The AR model of ground motion (acceleration component, for example) at a station can be expressed as

$$a_j[n]=\sum_{r=1}^{p}\alpha_j[r]a_j[n-r]+\varepsilon_j[n] \tag{2.1}$$

where a component of discrete ground motion record at a station j is denoted as $a_j[n]$, with $n=0,1,2,.....,N-1$, indicating the sample index. In total there are N samples, collectively referred to as a signal. The AR model parameters are represented by $\alpha_j[r]$, p is the model order indicating the number of past values used to predict the present value of the signal, and $\varepsilon_j[n]$ represents a one-step prediction error, considered to be sampled from a random (uncorrelated) white noise process with zero mean and variance σ_j^2. The model parameters are usually estimated using the least-squared minimization techniques. Variants of the ordinary least squares techniques are also available in the literature—such as, the Yule-Walker equations (Yule 1927, Walker 1931) and the Burg method (Burg 1968). The optimal order of the AR model can be selected by using well-defined statistical criteria, such as the Akaike's Final Prediction Error (FPE) criterion (Akaike 1970), Akaike's Information Criterion (AIC) (Akaike 1974), Parzen's Criterion of Autoregressive Transfer functions (CAT) (Parzen 1974). The AIC corresponding to a model order p
can be obtained as

$$AIC(p)=N\log\left(\sigma_j^2(p)\right) \tag{2.2}$$

where $\sigma_j^2(p)$ is the variance of the prediction error corresponding to an AR model of order p. The optimal model order minimizes the value of AIC. Once the model parameters are estimated, the parametric auto-spectral density is given as

$$\tilde{S}_{jj}(f)=\frac{\sigma_j^2\Delta t}{\left|1-\sum_{r=1}^{p}\alpha_j[r]\exp(-i2\pi rf\Delta t)\right|^2} \quad , |f|\le\frac{1}{2\Delta t} \tag{2.3}$$

where Δt represents the sampling interval. In the multivariate case, involving signals from m sensors with an equal number of samples uniformly spaced at Δt, seconds, the AR model becomes

$$\begin{Bmatrix} a_1[n] \\ a_2[n] \\ \vdots \\ a_m[n] \end{Bmatrix}=\sum_{r=1}^{p}\begin{bmatrix} \alpha_{11}[r] & \alpha_{12}[r] & \cdots & \alpha_{1m}[r] \\ \alpha_{21}[r] & \alpha_{22}[r] & \cdots & \alpha_{2m}[r] \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1}[r] & \alpha_{m2}[r] & \cdots & \alpha_{mm}[r] \end{bmatrix}\begin{Bmatrix} a_1[n-r] \\ a_2[n-r] \\ \vdots \\ a_m[n-r] \end{Bmatrix}+\begin{Bmatrix} \varepsilon_1[n] \\ \varepsilon_2[n] \\ \vdots \\ \varepsilon_m[n] \end{Bmatrix} \tag{2.4}$$

which can be compactly written as

$$\mathbf{a}[n]=\sum_{r=1}^{p}\boldsymbol{\alpha}[r]\mathbf{a}[n-r]+\boldsymbol{\varepsilon}[n] \tag{2.5}$$

where $\boldsymbol{\alpha}[r]$ is a $m\times m$ matrix of model parameters for each value of $r\in\{1,2,...,p\}$ and $\varepsilon_j; j\in\{1,2,....,m\}$ are jointly Gaussian zero-mean white noise processes with variances $\sigma_j^2, j\in\{1,2,...,m\}$. The optimal model order is selected by minimizing the AIC criteria for multivariate AR model, which is defined as:

$$AIC(p)=N\log\left(\det\mathbf{C}_\varepsilon(p)\right)+2m^2p \tag{2.6}$$

where $\mathbf{C}_\varepsilon(p)$ is the covariance matrix of ε_j corresponding to a model order p. The selected model can be viewed as a linear time-invariant system with the following transfer function matrix

$$\mathbf{H}(f)=\left[\mathbf{I}-\sum_{r=1}^{p}\boldsymbol{\alpha}[r]\exp(-i2\pi rf\Delta t)\right]^{-1},\ |f|\leq\frac{1}{2\Delta t} \tag{2.7}$$

with $\mathbf{I}$ as the identity matrix. The spectral matrix of the AR process can then be obtained as

$$\tilde{\mathbf{S}}(f)=\begin{bmatrix}\tilde{S}_{11}(f) & \tilde{S}_{12}(f) & \cdots & \tilde{S}_{1m}(f)\\ \tilde{S}_{21}(f) & \tilde{S}_{22}(f) & \cdots & \tilde{S}_{2m}(f)\\ \cdots & \vdots & \ddots & \vdots\\ \tilde{S}_{m1}(f) & \tilde{S}_{m2}(f) & \cdots & \tilde{S}_{mm}(f)\end{bmatrix}=\mathbf{H}(f)\mathbf{C}_{\varepsilon}\mathbf{H}^{*}(f)\Delta t \tag{2.8}$$

where the asterisk, *, denotes conjugate transpose of a matrix. The spectral matrix given by Eqn. 2.8 is Hermitian symmetric; its diagonal $\tilde{S}_{jj}(f)$ is the ASD (auto-spectral density) of the signal $a_j[n]$ and the off-diagonal terms $\tilde{S}_{ij}(f)$ are the CSD (cross-spectral density) between the signals $a_i[n]$ and $a_j[n]$. This approach of multivariate modelling allows the analyst to identify the spectral matrix that optimally describes all the observations of an array in the least-squares sense.

3 FREQUENCY-WAVENUMBER (F-K) SPECTRA

Frequency-wavenumber (f-k) spectra can be used to estimate backazimuth and apparent propagation velocity of impinging waves at the array. Useful information regarding the nature of the path followed by seismic waves from the rupture area to the array site can be inferred from such spectra. In addition, f-k analysis can also be used to understand the contribution of various wave components to the motion at array sensors. Several applications of f-k analysis can be found in the earthquake engineering and engineering seismology literature (see Zerva 2009 for a detailed reference list). Backazimuth (Baz) or direction of arrival (DOA) estimation along with apparent wave propagation velocity can be used to model wave passage effects in studying spatial variability of ground motion. The present study focuses on this application of f-k analysis. For a locally homogeneous and stationary random ground motion field, the f-k spectra can be defined as

$$F(\vec{\kappa},f)=|\Psi(\vec{\kappa},f)|^2 \tag{3.1}$$

where $\Psi(\vec{\kappa},\omega)$ is the time and space Fourier transforms of the ground motion field, f if the frequency, $\vec{\kappa}=\begin{bmatrix}\kappa_x & \kappa_y\end{bmatrix}^T$ is the wavenumber vector, the superscript T indicating transpose. The wavenumber is related to the slowness by the expression $\vec{s}=\vec{\kappa}/(2\pi f)$. The length of the slowness vector is equal to the inverse of the apparent propagation velocity c and its direction is the Baz, or opposite the direction of wave propagation. An estimate of the f-k spectra from array data with m sensors can be obtained as

$$\tilde{F}(\vec{s},f)=\frac{1}{m^2}\vec{u}^*(\vec{s})\tilde{\mathbf{S}}(f)\vec{u}(\vec{s}) \tag{3.2}$$

where $\tilde{\mathbf{S}}(f)$ is the spectral matrix; for sensor position vectors $\vec{r}_j$, $\vec{u}(\vec{s})=\exp\left[i2\pi f\vec{s}\cdot\vec{r}_j\right]$ is the beam-steering vector with complex number i. The slowness that maximizes the f-k spectra at each frequency gives, for a plane wave with the corresponding frequency, the propagation velocity $c=1/|\vec{s}|$. The direction of the corresponding slowness vector gives the Baz. In this sense c and Baz are frequency dependent in general. However, body waves are non-dispersive, and their slowness vector should be frequency independent. In such situations, the slowness vectors of different frequencies can be superimposed to produce the stacked slowness (SS) spectra, first introduced by

Spudich and Oppenheimer.

$$F(\vec{s}) = \sum_j \tilde{F}(\vec{s}, f_j) \tag{3.3}$$

4. EXAMPLE APPLICATION

In this section we demonstrate the application of the above described procedure to the strong motion array data obtained from the SMART1 array in Taiwan.

4.1. Data

The data being used are obtained from Event 5, which occurred on January 29, 1981, with a magnitude of $M_L = 6.3$, about 30 km from the array at a hypocentral depth of about 25 km. The configuration of the array is shown in Fig. 1. These data have been used by many researchers to study spatial variability of strong ground motion (Loh et al. 1982, Zerva 1986, Hao et al. 1989, Oliveira et al. 1991, to mention a few). Strong-motion data of the array were downloaded from the COSMOS virtual data centre.

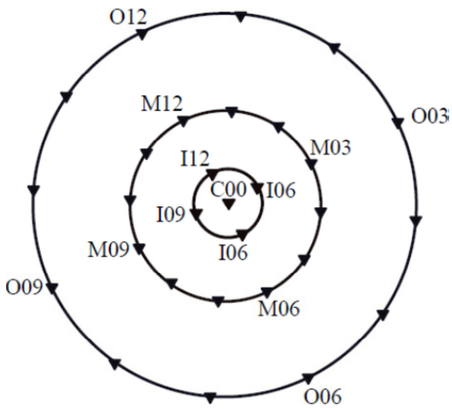

Figure 1. Configuration of the SMART 1 array showing the locations of the strong-motion stations (triangles). The central station is called C00; those at the inner ring, middle ring, and outer ring are numbered I01-I12, M01-M12, and O01-O12, respectively. The radii of the rings are 200m, 1km, and 2km.

4.2. Pre-processing and spectral analysis

The data used for illustration purposes are the horizontal components of the acceleration time series recorded at central, inner ring and middle ring stations. The recorded acceleration is numerically integrated to compute velocity time series, which are used in subsequent analysis. The S-wave window of 5.12 seconds (512 samples at 100 samples per second) is considered. If the time of first sample at station O12 is considered as the origin of time, the selected time window corresponds to 7.00 to 12.12 s. The signals are de-trended and tapered with a double cosine (Tukey) window, with taper durations of 15% of signal duration at both ends. AR parameters are computed by using least squares minimization techniques. Model order selection is based on AIC criteria, resulting in an optimal model order of 6. F-k spectra are computed in the frequency range of 0.2 to 20 Hz.

4.3. Results and conclusions

Figure 2 and 3 show the F-k spectra as a function of slowness in the two horizontal directions for the NS and EW components of recorded ground motions, respectively. The results corresponding to 6 different frequencies are shown. The spectral amplitudes are normalized by the maximum value for each frequency. The backazimuth is indicated by white line for each frequency. At 0.2 Hz an apparent velocity of 4.5 km/s is observed with and backazimuth of 153°, for the NS component. The results for frequencies up to 3 Hz are roughly consistent with that at 0.2 Hz. At 5 Hz, several spurious peaks are observed, with a considerably low apparent velocity. Similar observations can be made for the EW component results shown in Fig. 3. The spurious peaks may be contributed to spatial aliasing that can cause wavenumber shifts inversely dependent on the frequency (Spudich and Oppenheimer 1986). The high frequency motion thus seems to be controlled by scattered energy. These effects can be reduced, to a certain extent, by creating SS spectra.

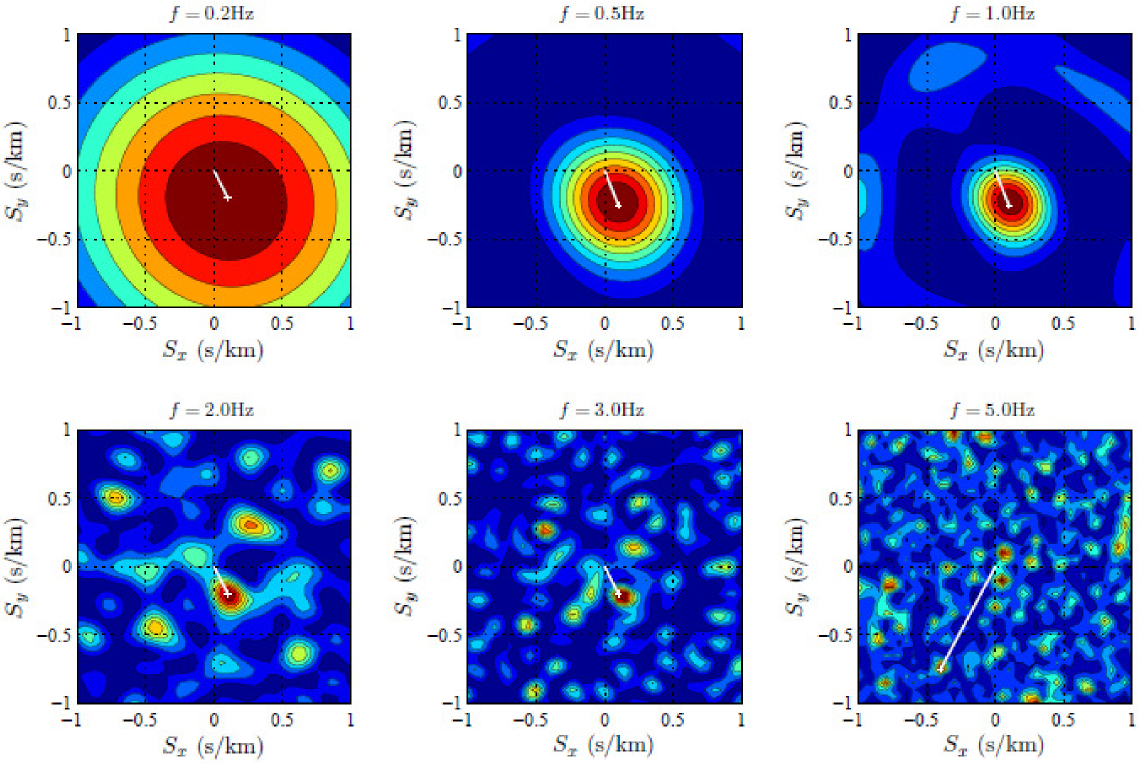

Figure 2. Slowness spectra computed from the NS component for 6 different frequencies as indicated in the title of the subplots. S_x and S_y indicate slowness in the in the EW and NS directions respectively. The spectral values from Eqn. 3.2 are normalized by the largest value. The location of peak is marked with a white cross, while the white line indicates the backazimuth direction.

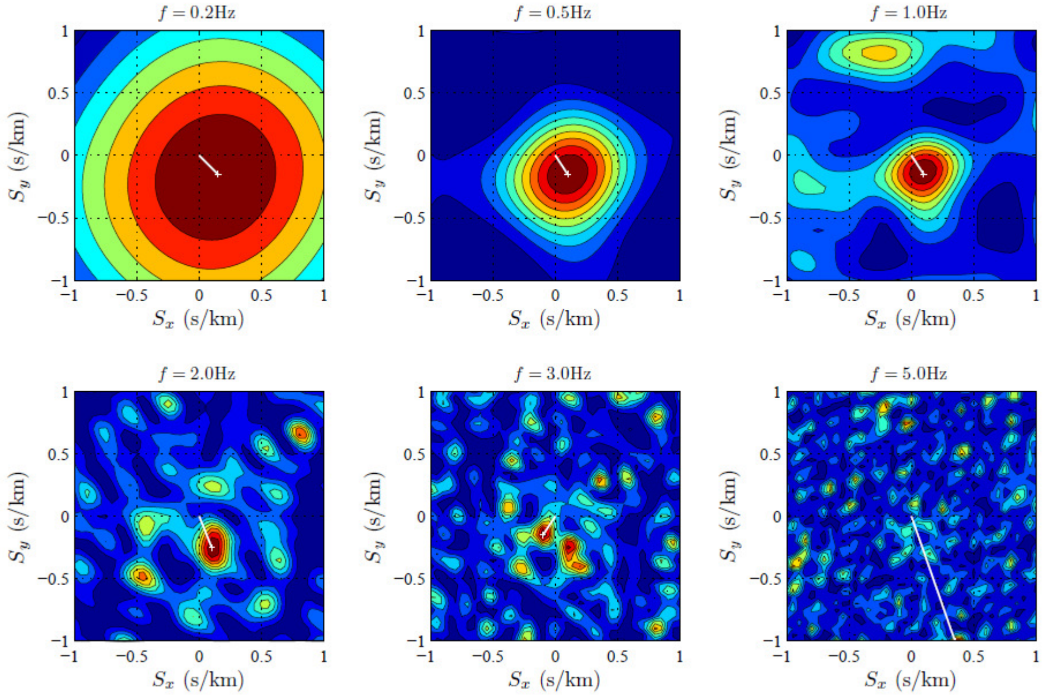

Figure 3. Same as in Fig. 2 above but for EW components of motion.

The stacked slowness spectra for the NS and EW components are shown in Fig. 4. The advantages of slowness stacking are clearly visible in the figure in terms of higher resolution, and reduced effects of scattered high frequency energy. For both components the identified slowness corresponds to 0.1 s/km and -0.2 s/km in the E and N directions, respectively. This results in apparent propagation velocity of 4.5 km/s, and a backazimuth of N153°E. These results are in perfect agreement with those reported in the literature (Zerva 2009). The estimated backazimuth is also consistent with the source-site geometry. Furthermore, it is noteworthy that the resolution of the slowness spectra is higher than that can be obtained by using conventional methods applying windowed periodograms (see Zerva 2009 for results obtained by conventional methods). Zerva (2009) also presents stacked slowness spectra obtained from the multiple signal characterization (MUSIC) method. In terms of resolution and accuracy, our results are comparable to the results obtained by MUSIC method. We note here that the application of MUSIC method to strong motion array data requires, to a certain degree, subjective decisions from the analyst in separating the signal and noise subspace, and also in sub array spectral averaging modification (see Goldstein and Archuleta 1999, and Bokelmann and Baisch 1991, for more details). In addition, the amplitudes of signals can only be approximately recovered from MUSIC estimates of f-k spectra, whereas the methods presented herein directly provides an estimate of both wavenumber and amplitude of the analyzed signals. This leads us to the conclusion that AR model based spectral estimation is potentially superior than other commonly employed methods, and requires more attention in spectral analysis of strong motion array data.

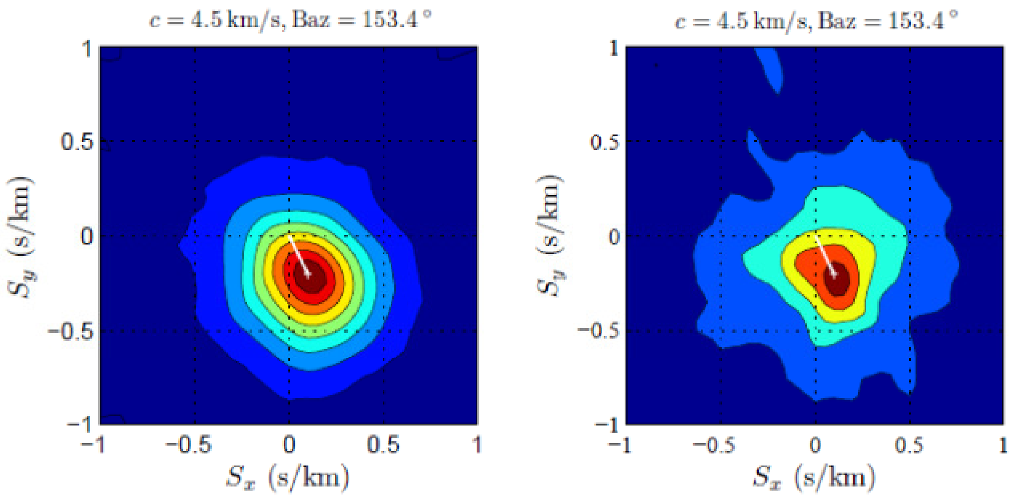

Figure 4. Stacked slowness spectra for NS and EW components on the left and right panels, respectively.

AKCNOWLEDGEMENT

This work was supported by research grants from the Ludvig Storr Cultural and Research Foundation and Landsvirkjun's Energy Research Fund. Furthermore, we acknowledge the support from the University of Iceland Research Fund

REFERENCES

Akaike, H. (1970) Statistical predictor identification. *Ann. Inst. Statist. Math* **22**,203-217

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Trans. on Automatic Control* **AC:19**, 716-722

Beamish, N. and Priestley. M. B. (1981) A study of Autoregressive and Window Spectral Estimation. *Journal of the Royal Statistical Society, Applied Statistics* **30:1**,41-58

Bokelmann. G. H. R. and Baisch, S. (1999) Nature of narrow band signals at 2.083 Hz. *Bulletin of the Seismological Society of America.* **89**,156-164

Box, G., Jenkins, G. M. and Reinsel, G. C. (1994) Time Series Analysis: Forecasting and control, third edition, Prentice-Hall

Burg, J. P. (1968) A new analysis technique for time series data. *NATO Advanced Study Institute on Signal Processing, Enschede, Netherlands, August 1968*.

Broersen, P. M. T. (2006) Automatic autocorrelation and spectral analysis, Springer-Verlag London Limited.

Goldstein, P. and Archuleta, R.J. (1991) Deterministic frequency wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array : Data Analysis. *Journal of Geophysical Research* **96**,6187-6198

Hao, H., Oliveira, C. S. and Penzien, J. (1989) Multiple station ground motion processing and simulation based on SMART1 array data. *Nuclear Engineering and Design* **111**,293-310

Hindy, A. and Novak, M. (1980) Pipeline response to random ground motion. *Journal of the Engineering Mechanics Division, ASCE* **106:2**,339–360

Kay, S. M. and Marple, S. L. (1981) Spectrum analysis: a modern perspective. *Proc IEEE* **69**,1380-1419

Kozin, F. (1988) Autoregressive moving average models of earthquake records. *Probabilistic Engineering Mechanics* **3:2**, 58-63

Lenonard, M. and Kennett, B.L.N. (1999) Multi-Component autoregressive techniques for the analysis of seismograms. *Physics of the Earth and Planetray Interiors* **113**, 247-263

Ljung, L. (1999) System identification: Theory for the user, second edition. Prentice Hall, Upper Saddle River, NJ 07458

Loh, C. H., Penzien, J. and Tsai, Y. B. (1982) Engineering analysis of SMART1 array accelerograms. *Earthquake Engineering and Structural Dynamics* **10**,575-591

Newmark, N. M. and Rosenblueth, E. (1971) Fundamentals of Earthquake Engineering, Prentice Hall, Inc., Englewood Cliffs, NJ.

Oliveira, C. S., Hao, H. and Penzien, J. (1991) Ground motion modeling for multiple input structural analysis. *Structural safety* **10**:79-93

Parzen, E. (1974) Some recent advances in time series modelling. *IEEE Trans. on Automatic Control* **AC(19),**

723-730
Pavageau, M., Rey, C. and Elicer-Cortes, J-C. (2004) Potential benefit from the application of autoregressive spectral estimators in the analysis of homogeneous and isotropic turbulence. *Experiments in fluids* **36**, 847-859
Spudich, P. and Oppenheimer, D. (1986) Dense seismograph array observations of earthquake rupture dynamics, in *Earthquake Source Mechanics, Geophysical Monograph 37, S. Das, J. Boatwright and C.H. Scholz editors, American Geophysical Union, Washington, D.C.*
Thostehim, D. (1975) Some autoregressive models for short-period seismic noise. *Bulletin of the Seismological Society of America* **65:3**, 677-691
Walker, G. (1931) On periodicity in series of related terms. *Proceedings of the Royal Society of London* **A:131**,518-532
Yule, G. U. (1927) On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical transactions of the Royal Society of London* **A(226)**,267-298
Zerva, A. (1986) Stochastic differential ground motion and structural response. Ph.D. Thesis, Department of Civil Engineering, University of Illinois at Urbana Champaign, Urbana IL
Zerva, A. (2009) Spatial variation of seismic ground motions, Modelling and engineering applications. CRC Press, Taylor and Francis Group, LLC.