GEOPHYSICAL RESEARCH LETTERS, VOL. 22, NO. 4, PAGES 513-516, FEBRUARY 15, 1995

Reshaping spectrum estimates by removing periodic noise: Application to seismic spectral ratios

Jonathan M. Lees

Department of Geology and Geophysics, Yale University, New Haven, Connecticut

Abstract. An automated method for removing line spectrum elements embedded in colored spectra is presented. Since smooth spectrum estimates are desired, line spectra tend to smear out over an effective smoothing window. This introduces a bias in spectrum estimation that seriously degrades signal-to-noise ratios, spectral deconvolution or any other operation where spectrum shape is important. Multi-taper analysis provides a simple algorithmic solution including a method of determining where spectral peaks are both significant with high power. The method is completely general, and examples include estimation of signal-to-noise ratio at the 1990 high frequency array, Pinyon Flat, CA. A comparison of noise spectra line segments and signal spectra line spectra reveals similarities associated with instrument noise and shallow resonances stimulated by incoming seismic signals. Identification and removal of resonances can provide a better means of estimating background noise spectra for modeling earthquake source spectra and path effects associated with attenuation.

Introduction

There are many situations in geophysical signal analysis where we are presented with time series that include a continuous background spectrum punctuated with periodic signals. A classic example in seismology is the "white" noise spectrum of pre-event recordings contaminated by 60 Hz power line noise. In most data sets the location of spectral lines are not known *a priori*, and often spectral lines are the very target of the research. In other situations spectral lines are viewed as noise and the continuum colored spectrum is the object of analysis. If the periodic part of the signal is known *a priori*, it can often be removed explicitly, prior to analysis, or it may be filtered from the trace, on the fly, in the course of analysis. Filtering, though, introduces additional distortions to spectra due to finite filter roll-off and limited bandwidth resolution. Furthermore, presence of a strong line spectrum representing noise does not mean that the underlying background signal has a large hole where noise has been filtered. Rather, we desire a background spectrum estimate which replaces the line spectrum with a new spectrum estimate based on the underlying properties of the data.

In a seminal paper on multi-taper spectrum analysis, Thomson [1982] outlined a new method for precise determination of spectral lines using a statistical test which follows naturally from the multi-taper analysis. He further described a method for removing the periodic signals and "reshaping" the spectrum in the vicinity of the lines to achieve a better approximation to the (assumed) continuum background spectrum. Thomson's methods have been applied before to geophysical and geologic time series in several papers[*Park et al.*, 1987a; *Park et al.*, 1987b; *Park et al.*, 1987c; *Vernon et al.*, 1991]. In this paper I present examples of Thomson's method applied to seismic signals and noise at Pinyon Flat Observatory, California. I develop and present an automated procedure for identifying periodic signals buried in otherwise white noise or colored background spectra.

Paper number 94GL03221
0094-8534/95/94GL-03221$03.00

Multi-taper analysis

Multi-taper spectrum analysis computes a spectrum by combining a set of independent, orthogonal spectrum estimates. These are calculated in three steps: first a set of orthogonal (Slepian) tapers are calculated by solving an eigenvector problem. Second the tapers are sequentially applied to the data window followed by Fourier transformation. Fourier transforms of the Slepian tapered data are called eigenspectra. Eigenspectra are finally combined (averaged) to form the final spectrum estimate. The way the spectra are combined depends on the situation at hand. In this paper we have used adaptive algorithms outlined in Thomson[1982] involving iteratively solving a complex regression problem. As with all spectrum analysis, there is a compromise between resolution and variance. One advantage of the multi-taper scheme is that it provides an optimal compromise given the assumptions behind the creation of the tapers (the eigenvalue problem) and the methods of averaging the eigenspectra. Another virtue of the multi-taper approach is that it reduces bias associated with finite windowing of an otherwise infinite data stream.

The main advantage of the multi-taper approach, however, lies in the statistical properties following from the theory. Thomson outlined an algorithm which allows us to identify and isolate periodic line spectra buried in a colored background spectrum. His approach is based on an F-test which compares the variance of an assumed locally white noise background to the variance explained by a periodic signal. Since we are interested in smoothed spectrum estimates, a monochromatic spike in the spectrum will be smeared out over an effective averaging window, $F_W = \eta / (N * \Delta t)$, where η is the parameter controlling the order of the Slepian tapers, N is the number of samples and Δt is the sampling rate. In this study $\eta = 3.0$ and 5 Slepian tapers were used, although other combinations of these parameters produced nearly identical results. Once the signal has been identified, it can be removed, and the biased spectrum (due to smoothing) near the line element can be reshaped. The method is explained in Thomson [1982], and Percival and Walden [1993] (equations 104, 499a-c, 500) and computer programs are provided by Lees and Park [1994]. Usually these methods are applied to single time series where automation is not necessary and careful examination of the spectra and periodic elements is important. In this paper we provide an automated way of searching for and reshaping distorted spectra due to periodic noise which contaminate estimates of signal-to-noise ratios in seismic analysis. These techniques might be applied to data sets where extensive study of the individual spectra is prohibitive. The automated program could be applied to large data sets or included within a set of standard processing procedures applied in real time.

To insure that the statistical F-test indicates a line spectrum above background, one must examine the spectrum itself for high power at those frequencies. We are interested in isolating frequencies that have local (as opposed to global) high spectral power. Since the spectrum may include a linear trend, or other red noise background, we attempt to remove this trend prior to determining if a particular peak in the spectrum has power above the background. In this study the spectrum (not the time series) was flattened by simply removing the mean and linear trend. The flattening is somewhat arbitrary and depends on the expected background spectrum process. For earthquake spectra presented below a reasonable alternative may include removing an additional quadratic factor. If the background spectrum has some associated a priori model, like a red noise auto-regressive spectrum, then it may be helpful to subtract out such a model prior to determining if the spectrum has local power above the background. I have found that removing a robust, median smoothed spectrum estimate often works well. With the "long wavelength" features of the spectrum removed, the "short wavelength", or local, spectral features can be compared. The logarithm of the flattened spectrum is then examined for large values, typically 5 standard deviations from the mean (a conservative cut off range). If the signal power is large at those points where the F-test also (independently) indicates high significance, that frequency is designated for removal (or reshaping).

Reshaping involves replacing the multi-taper estimate with a locally less biased estimate in a window centered at the frequency which passed the F-test and power tests outlined above (Percival and Walden [1993], equation 500). This new spectrum is less biased over the smoothing window and provides a better (pointwise) spectrum estimate for use in deconvolution and spectral division. If the periodic signal itself is relevant to the analysis, it can be replaced at its specific frequency as a line element. This analysis is carried out in Thomson's 1982 paper and examples using tidal data are presented in Percival and Walden [1993].

Examples

Modeling the shape of seismic spectra is becoming an important tool for estimating source parameters and path effects (attenuation) in three-dimensional seismic inversions[*Lees and Lindley*, 1994]. Unbiased spectrum estimates are particularly desirable when using spectra for calculating spectral ratios. Examples include estimating signal-to-noise ratios, calculating transfer functions between borehole and surface recordings, dividing spectra to calculate receiver functions[*Owens et al.*, 1988] and estimating site and path effects by eliminating seismic source effects via spectral deconvolution[*Evans and Zucca*, 1988]. If the shape of the spectral ratio is later used for modeling some specific process (e.g. displacement spectrum in seismic attenuation) then introduction of biases may severely alter estimation of model parameters. In this paper we show how biases can be eliminated in estimating signal-to-noise ratios.

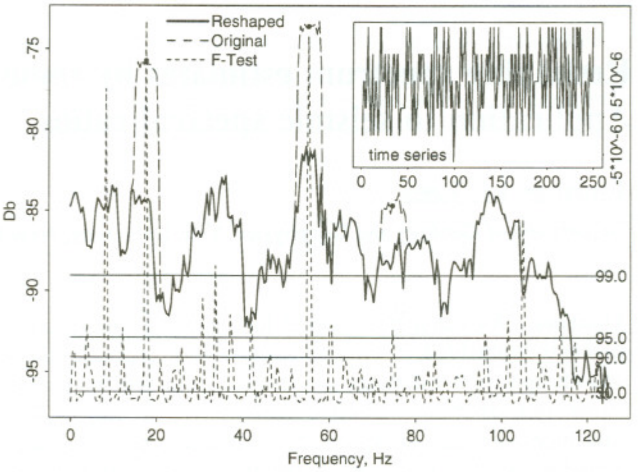

Figure 1: Time series (inset), original (long dash) and reshaped (bold) spectra for pre-event noise. F-test (short dash) isolates frequencies which surpass confidence levels taken from published tables. Bullet marks are points in the spectrum designated for reshaping.

As an example consider a (pre-event) time series (Figure 1) taken from the data set described below. The F-test has been scaled to overlay the spectrum, so visual correlation in bands of high F-test and high power are evident. (This example was chosen, in part, for its low level of pre-event noise, and thus quantization noise may introduce spurious noise. Other examples are not as strongly affected.) Three frequencies are designated for reshaping: 17.5, 55 and 75 Hz (marked by bullets). The significant F-test at 12 Hz was not reshaped because local power in the spectrum was not considered strong enough to merit reshaping. When computing signal-to-noise ratios, we should use the new, reshaped spectrum for deconvolution. Figure 2 shows signal (post-event time series) and noise spectra, each plotted with their original and reshaped spectra. Notice that periodic line spectra in the noise and signal time series do not always correspond.

In 1990 a PASSCAL working group deployed 60 seismic stations at Pinyon Flat Observatory, California (Figure 3)

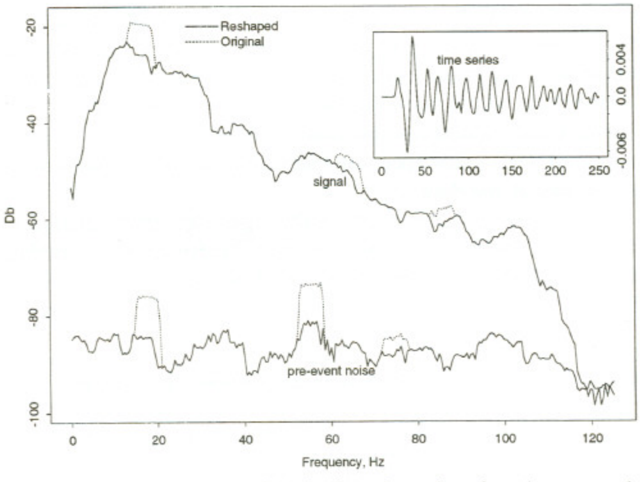

Figure 2: Post-event signal (inset) and noise (pre-event) spectra.

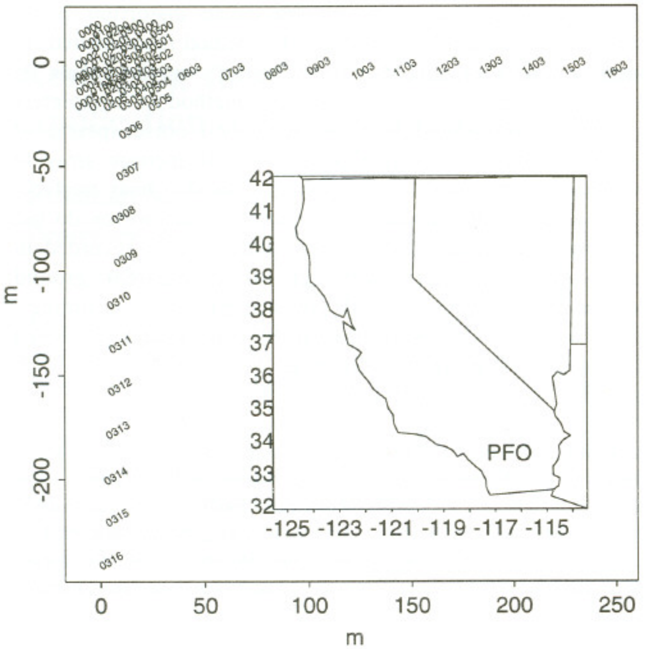

Figure 3: Configuration of Pinyon Flat High Frequency Array. Inset shows location in California. Surface deployment included a square with two arms trending south and east, and two borehole seismometers installed at 150 and 250 meters depth.

spatially distributed over 259 m. The tight array consisted of 2 borehole seismometers and 58 surface seismometers laid out in a square with two orthogonal arms projecting north-south and east-west [*Owens et al.*, 1991]. Each station utilized a three-component seismometer (L-22) recording at 0.004 ms. Seismometers were left to record earthquakes for about 6 weeks and tens of events, mostly located southwest in the San Jacinto Fault zone, were recorded on nearly all the stations at the observatory. A typical sample event (90.134.05.05.20, station x03y03) is located approximately 20 km southwest of the PFO observatory (Figure 3).

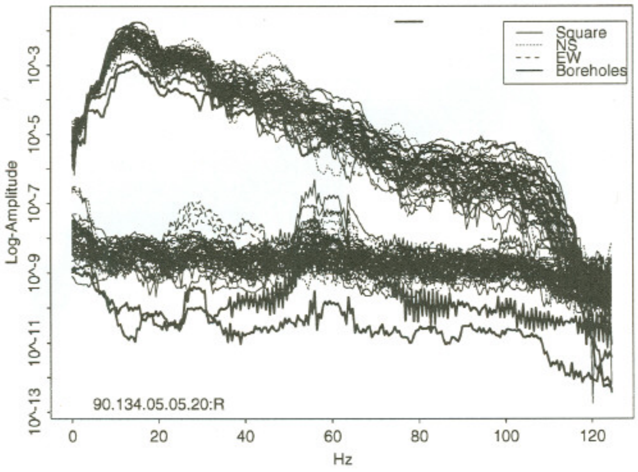

Figure 5: Reshaped spectra of Pinyon Flat data. Reshaping has removed most of the bias associated with periodic elements, like at 60 Hz.

Figure 4 shows a plot of original spectrum estimates for noise and signals for one event recorded at all 60 stations shown in Figure 3. It is known that about 20% of L22's have spurious resonances between 22-29 Hz [*Menke et al.*, 1991]. Here there are several stations which further exhibit low frequency resonances (typically between 15-20 Hz). If the resonance and instrument noise is monochromatic, the smoothed spectrum will be biased over the smoothing window and estimates of the signal-to-noise ratio will be biased. This represents a strong candidate for spectral reshaping, as division of the noise into the signal may fluctuate rapidly in bands that have large bias. Figure 5 shows the reshaped spectra for the 60 stations. There is still some left over bias in some of the noise spectra near 60 Hz although the reshaping reduced the bias by several orders of magnitude. Note that the noise at 17.5 Hz was completely eliminated. The signal spectra were also reshaped using the same algorithm, although the effects are not as spectacular because periodic noise has less power relative to the signal as compared to the periodic signals embedded in the (assumed) "white" spectrum of the pre-event time series.

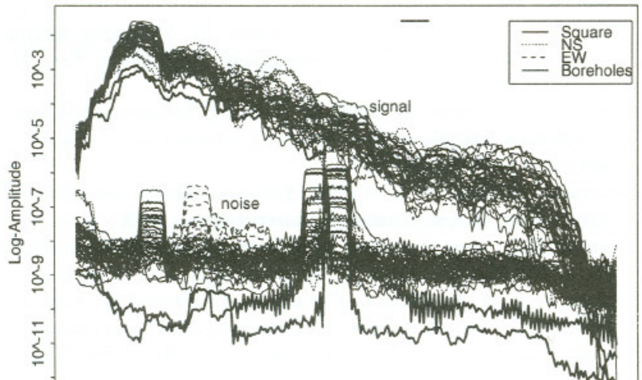

Figure 4: Post and pre-event spectra (1 second windows) for event 90.134.05.05.20. Lines are coded so that seismometers in the square can be distinguished from east-west (EW) and NS lines. Boreholes are plotted in bold.

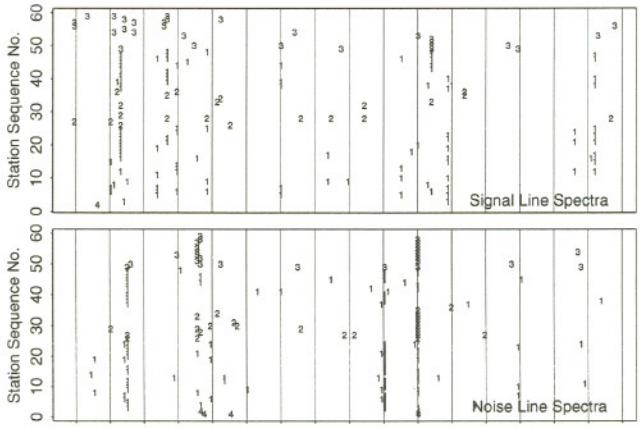

Figure 6: Compilation of the periodic spectral peaks removed from spectra in Figure 5. Upper plot is for the post-event spectra and lower plot is for the pre-event spectra. For each reshaped line a central frequency is plotted with 1=square, 2=EW-arm, 3=NS-arm, 4=Boreholes.

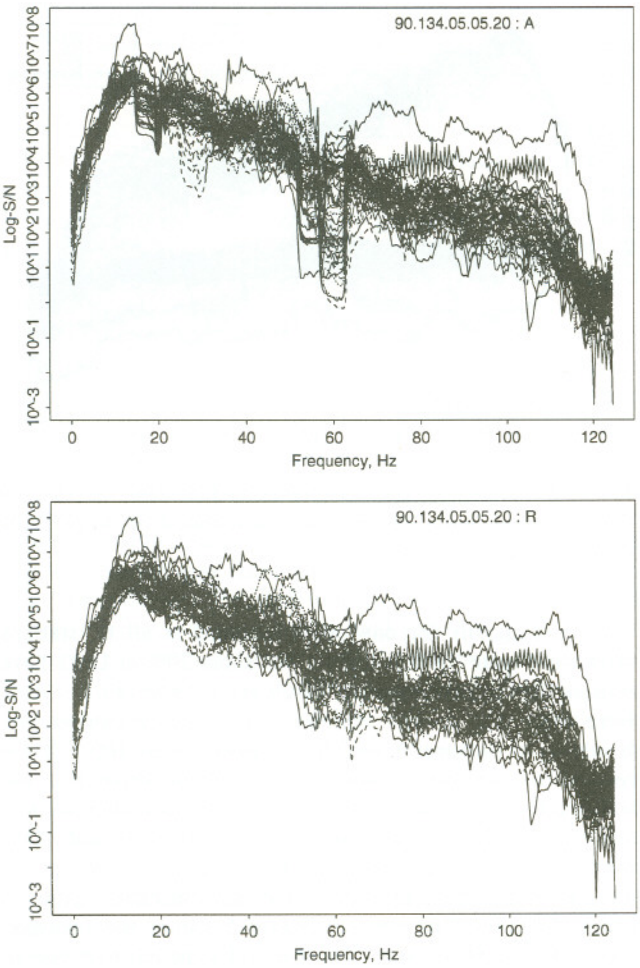

Figure 7: Estimates of signal-to-noise in the original (A) and reshaped spectra (R). Logarithm ratios show extreme bias near 55-60 Hz and apparent bias just below 20 Hz. These biases have been, for the most part removed by the reshaped spectra below.

If we tabulate the periodic line spectra and plot them together we can see bands that have consistently strong periodic signals in both the post and pre-event time series (Figure 6). The pre-event series appear to share line spectra at 17.5, 28, 55, and 60 Hz. The post event signals appear to have line spectra at 17 Hz, several consistent peaks between 20-30 Hz and some regular features above 60 Hz, although the signal line spectra are not as consistent as the noise line spectra. A detailed discussion of the nature of the spatial clustering of these signals is beyond the scope of this paper, and is presented in a companion paper.

Results of signal-to-noise ratio estimates are determined for original and reshaped spectra (Figure 7). Improvements in signal-to-noise ratios are pronounced where periodic noise at 55-60 Hz signals are dominant in the noise spectra. A more subtle effect is evident just below 20 Hz and around 28 Hz. In this example reshaped spectral estimates of signal-to-noise ratio are clearly superior to original, non-reshaped estimates.

Conclusion

I have presented and outlined a method for attaining a less biased smoothed estimate of time series spectra when the presence of periodic line spectra cause distortion in an otherwise colored noise process. The method is automated, so large amounts of data can be "cleaned" without hands on analysis of each signal. While the method is completely general, it is applied here to signal and noise spectra of seismic signals in a high frequency small aperture array to investigate the nature of resonance in the near surface. Seismic noise stimulated by incoming seismic waves do not necessarily correlate with ambient ground noise recorded in pre-event time series, although some apparent ground resonance does appear in both pre- and post-event recordings. Signal-to-noise ratio estimates can be considerably improved by using reshaping to eliminate biases associated with periodic noise in seismic signals.

Acknowledgments. PFO data was acquired by P. Anderson, J. Babcock, D. McNamara, T. Owens, G. Pavlis, and F. Vernon. The author thanks Jeffrey Park, Mike Mann, Gary Pavlis and two anonymous referees for critical comments. This work was partially supported by NSF NEHRP grant EAR-9011441, the donors of The Petroleum Research Fund, PRF 26595-G2, administered by the American Chemical Society.

References

Evans, J. R., and J. J. Zucca, Active high-resolution seismic tomography of compressional wave velocity and attenuation structure at Medicine Lake volcano, northern California Cascade Range, *J. Geophys. Res.*, *93*, 15016-15036, 1988.

Lees, J. M., and G. T. Lindley, Three-dimensional attenuation tomography at Loma Prieta: Inverting t* for Q, *J. Geophys. Res.*, *99*, 6843-6863, 1994.

Lees, J. M., and J. Park, Multi-taper spectral analysis: A stand-alone C-subroutine, in press *Computers & Geology*, 1994.

Menke, W., L. Shengold, G. Hongsheng, H. Ge and A. Lerner-Lam, Performance of the short period geophones of the IRIS/PASSCAL array, *Bull. Seismol. Soc. Am.*, *81*(1), 232-242, 1991.

Owens, T. J., P. Anderson and D. E. McNamara, The 1990 Pinyon Flat high frequency array experiment, *PASSCAL Data Report* , 91-002, 11 pp., 1991.

Owens, T. J., R. S. Crosson and M. A. Hendrickson, Constraints on the subduction geometry beneath western Washington from broadband teleseismic waveform modeling, *Bull. Seis. Soc. Am.*, *78*, 1319-1334, 1988.

Park, J., C. R. Lindberg and D. J. Thomson, Multiple-taper spectral analsyis of terrestrial free oscillations: Part 1, *Geophys. J. R. Astr. Soc.*, *91*, 755-794, 1987a.

Park, J., C. R. Lindberg and F. L. Vernon, III, Multitaper spectral analysis of high-frequency seismograms, *J. Geophys. Res.*, *92*, 12,675-12,684, 1987b.

Park, J., F. L. Vernon, III and C. R. Lindberg, Frequency dependent polarization analysis of high-frequency seismograms, *J. Geophys. Res.*, *92*, 12,664-12,674, 1987c.

Percival, D. B., and A. T. Walden, *Spectral Analysis for Physical Applications*, 583 pp., Cambridge University Press, 1993.

Thomson, D. J., Spectral estimation and harmonic analysis, *Proc. IEEE*, *70*, 1055-1096, 1982.

Vernon, F. L., J. Fletcher, L. Carroll, A. Chave and E. Sembera, Coherence of seismic body waves from local events as measured by a small-aperture array, *J. Geophys. Res.*, *96*, 11,981-11,996, 1991.

J. M. Lees, Department of Geology and Geophysics, Yale University, P.O. 208109, New Haven, CT 06520-8109.
(e-mail: lees@lamb.geology.yale.edu)

(Received: September 9, 1994; accepted November 7, 1994)