Bulletin of the Seismological Society of America, Vol. 87, No. 4, pp. 961–970, August 1997

A Bayesian Nonlinear Inversion of Seismic Body-Wave Attenuation Factors

by S. Gao[1]

Abstract It is a well-known fact that the uncertainties in measuring relative attenuation factors within a local or regional seismic network are usually high, due to noise of different kinds and unrealistic assumptions. Numerical experiments using nine synthetic seismograms, created using t^* values ranging from 0.1 to 0.9 sec, reveal that the commonly used spectral ratio method is strongly affected by the selection of data processing parameters such as width of the spectral smoothing window, reference station, and so on. The numerical experiments demonstrate that a Bayesian nonlinear inversion approach that directly matches the spectra is better at finding the correct parameters used to generate the synthetic seismograms. The Bayesian inversion approach uses *a priori* information to simultaneously search for the t^* values, the common spectrum for all the records from an event, and the near-receiver amplification factors by using all the recordings from an event. When z, the ratio of Gaussian noise to signal, $\leqq 0.1$, the spectral ratio and Bayesian methods yield similar results with mean t^* measurement errors <0.05 sec. For $0.1 < z \leqq 0.8$, the mean errors of the spectral ratio method are larger than 0.1 sec and in some cases as large as 0.6 sec, while those of the Bayesian method are less than 0.08 sec. Frequency-independent t^* and near-receiver amplification factors are assumed. A multi-step procedure is proposed to reject records with a large misfit.

Introduction

Measurements of seismic wave attenuation factors, quantified by t^* or travel time over Q, provide important information about the physical state of the earth (e.g., Knopoff, 1964; Anderson, 1967; Solomon and Toksöz, 1970; Jackson and Anderson, 1970; Der *et al.*, 1975; Taylor *et al.*, 1986). However, the uncertainties of measurements are usually high, due to different kinds of noise and unrealistic assumptions. An important criterion for any practical method is the stability of the measurements with regard to data processing parameters. In this study, synthetic data are used to compare two methods for finding the t^* values used in creating the data. It was found that a Bayesian nonlinear inversion method is better than the popular spectral ratio method.

The amplitude spectrum of event k recorded by station i, $A_{ik}(f)$ can be written as (e.g., Teng, 1968)

$$A_{ik}(f) = S_k(f)G_{ik}(f)R_{ik}(f)I_i(f), \tag{1}$$

where $S_k(f)$ is the spectrum of source waveform $S(t)$, $G_{ik}(f)$ is that of a Green's function $G(t)$, $R_{ik}(f)$ is that of the near-receiver effects, and $I_i(f)$ is that of the instrument response. The spectrum of the Green's function can be written as

$$G_{ik}(f) = \exp[-\pi f t^*_{ik}(f)], \tag{2}$$

where t^*_{ik} is the attenuation factor that is defined as the ratio of travel time to the effective mean quality factor Q; that is,

$$t^*_{ik}(f) = \int_{\text{path}} \frac{ds}{Q(s,f)V(s,f)}, \tag{3}$$

where $V(s, f)$ is the instantaneous velocity of body waves. Both V and Q are in general a function of location as well as frequency.

The main causes of high measurement uncertainties include the following: (1) noise caused by natural and cultural activities; (2) scattering caused by velocity fluctuations (Richards and Menke, 1983); (3) Near-receiver topographic effects (e.g., Vidale *et al.*, 1991; Frankel and Leith, 1992); (4) focusing/defocusing of energy from sedimentary lens structures (e.g., Gao *et al.*, 1996); (5) frequency dependence of $t^*_{ik}(f)$ when a frequency-independent approach such as the spectral ratio method is used (e.g., Anderson and Given, 1982); (6) uncertainties in the determination of source-related effects such as focal mechanisms and radiation patterns; and (7) interference from other arrivals that can be avoided by using phases that are well separated from other possible arrivals, and by careful selection of time windows for the computation.

[1]Present address: Department of Terrestrial Magnetism, Carnegie Institution of Washington.

A large portion of the noise introduced in the above processes can be characterized as random. In this study, we use Gaussian-noise-added synthetic data to test the capability of two approaches, the commonly used spectral ratio method and a less commonly used Bayesian nonlinear inversion approach in finding the correct t^*.

Data Generation

The nine synthetic seismograms (Fig. 1) are generated by using the inverse Fourier transforms of the spectra computed using equations (1) and (2) by letting $R_i(f) = 1$ and $I_i(f) = 1$. The source spectrum, $S(f)$, is taken as the Fourier transform of a typical short-period teleseismic body-wave record. Note that because only one event is involved, the event subscript k is omitted. Each record is assigned a different attenuation factor ranging from 0.1 to 0.9 sec using $t_i^* = i \times 0.1$, where i is the record number. Also shown in Figure 1 are the peak amplitudes of the resulting records. Gaussian noise of different levels is added to the synthetic seismograms; that is,

$$W(t) = W_0(t) + z \times N(t), \tag{4}$$

where $W(t)$ is the synthetic data, $W_0(t)$ is the signal that is the convolution of $S(t)$ and $G(t)$ in equation (1), $N(t)$ is a series of random numbers with a Gaussian distribution of zero mean and unit variance, and z is the relative noise level ranging from 0 to 0.8 times the absolute peak value of $W_0(t)$. The Gaussian noise was generated using a random number generator (Press *et al.*, 1992). Care has been taken to ensure that the noise on different records are independent. The sampling rate of the synthetic records is 10 samples per second. Examples of the synthetic records and their Fourier spectra are shown in Figure 2, where $z = 0.3$ is used.

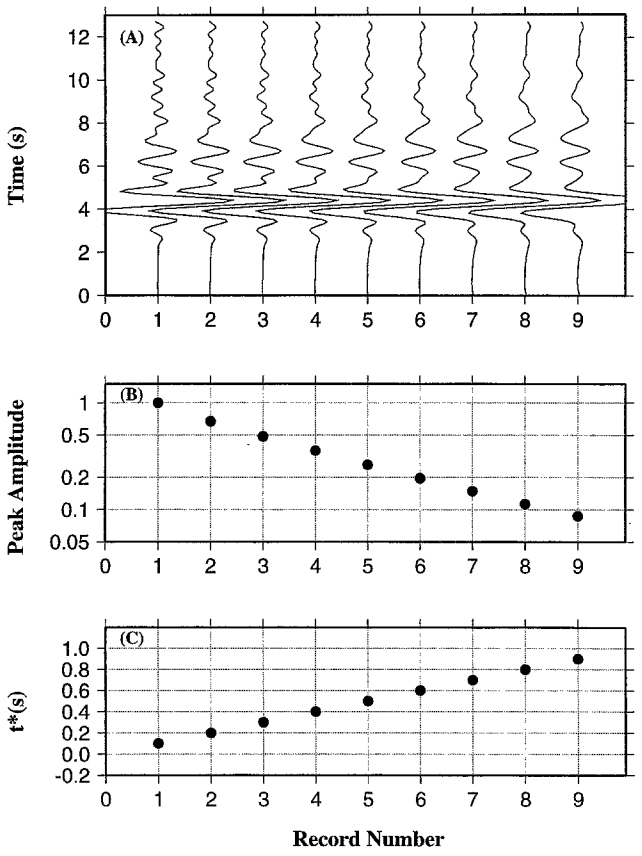


Figure 1. (a) Noise-free synthetic seismograms created using different t^* values. The seismograms were normalized by the peak amplitude of each trace, which is shown in (b). The t^* values are shown in (c).

Spectral Ratio Method

Most seismic body-wave attenuation studies have used the spectral ratio method (e.g., Teng, 1968; Solomon and Toksöz, 1970; Der and McElfresh, 1976). In this method, the spectral ratio between two stations, station i and the reference station j, is used to determine δt_{ij}^*, the relative attenuation factor between the two stations.

From equation (1), after instrument correction and under the assumption that

$$\ln \frac{R_i(f)}{R_j(f)} = C, \tag{5}$$

where C is a constant, the logarithm of the spectral ratio becomes a linear function of frequency, and the difference

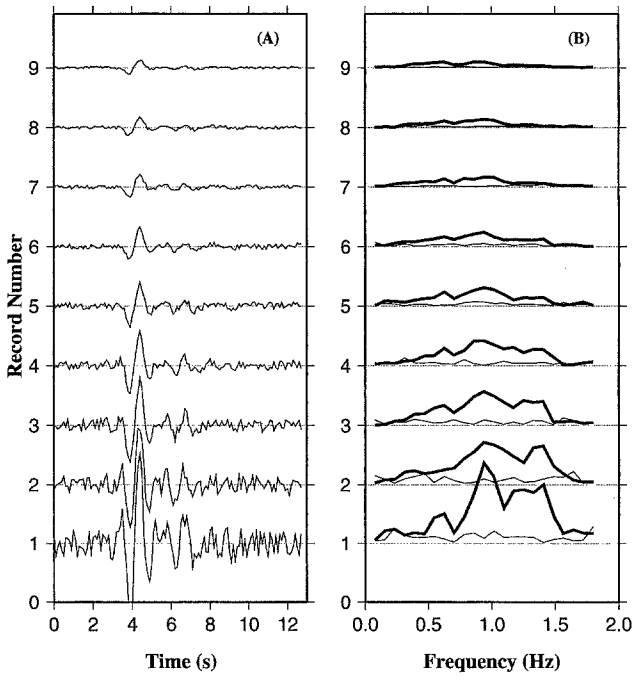


Figure 2. (a) Synthetic seismograms with noise/signal = 0.3. (b) Total spectra (thick lines) and noise spectra (thin lines) of the corresponding seismograms in (a).

in t^*, δt_{ij}^*, can be obtained by fitting the logarithm of the spectral ratio with a straight line; that is,

$$\ln \frac{A_i(f)}{A_j(f)} = C - \pi \delta t_{ij}^* f. \tag{6}$$

To calculate δt_{ij}^*, a time window of 12.8 sec was taken starting from 6.4 sec before the onset of the signal, and a noise section of the same length was taken before this time window. Both sections were tapered with a cosine taper. Fourier amplitude spectra were obtained from both sections and smoothed with a moving window of a chosen width. Estimation of signal spectra was obtained by subtracting the smoothed noise spectra from the total spectra. Data processing parameters include the following:

F_s: length of frequency window for smoothing.
T: length of time window for computing spectra. $T = 12.8$ sec is used. When the method is applied to real data sets, some care has to be taken to window out the arrival of interest so as to avoid interference of secondary arrivals.
T_0: time before the onset of actual signal. In this study we use $T_0 = T/2 = 6.4$ s. This will put the onset of the actual signal at the center of the tapering window.
SN: cutoff ratio between total amplitude and noise amplitude. Data points with a ratio smaller than SN are rejected. In this study, we use SN = 2.0.
bp_1, bp_2: bandwidth for spectral ratio. In this study, we use $bp_1 = 0.1$, $bp_2 = 2.0$ Hz. Therefore the number of spectral points is $(bp_2 - bp_1)/\delta f + 1 = 24$, where sampling rate in the frequency domain $\delta f = 1/(0.1 \times 128)$ Hz.
S_0: reference record.

Measured and fitted spectral ratios using low-noise data ($z = 0.1$) are shown in Figure 3a. The slopes of the fitted straight line divided by $-\pi$ are the t^* values relative to that of record 1, which is 0.1 sec. Another example with $z = 0.3$ is shown in Figure 3b, where about ⅓ of the data points were rejected, because the total amplitude is less than SN times the noise amplitude at these frequencies on either of the two records involved in the calculation of spectral ratio. By varying some of the data processing parameters mentioned above, we found that the spectral ratio method is not stable for data with high z. For instance, in Figure 3c, we used exactly the same parameters as those used in Figure 3b, except that record 5 instead of record 1 was used as the reference record. The resulting t^* values relative to record 1 are very different between the two tests, with differences as large as 0.1 sec. For noisier data, the difference could be as large as 0.3 sec. Different F_s values also result in significantly different results, as shown in Figures 3d and 3b, where $F_s = 4$ points and 2 points, respectively.

The mean measurement errors for different z values are shown in Figure 4. When $z \leqq 0.1$, the errors of the measurements are within 0.05 sec. For $0.1 < z \leqq 0.8$, the mean errors are larger than 0.1 sec and in some cases as large as 0.6 sec. Figure 4a was obtained using $F_s = 2$ points and record 1 as the reference record. Results shown in Figure 4b were obtained using $F_s = 2$ but record 5 as the reference record. Figure 4c was obtained using $F_s = 4$ points and record 1 as the reference record. Comparison of Figures 4a, 4b, and 4c indicates that the selection of reference record and the length of smoothing windows affect the measurements strongly.

Common Spectrum Method

The spectral ratio method uses a single reference record to determine δt^*. As demonstrated above, results from such a method are often unstable. They are strongly affected by the spectrum of the reference record. To overcome such a problem, Halderman and Davis (1991) used a nonlinear inversion procedure, called the common spectrum (CS) method. They used the inversion method in Bevington (1969) to directly match the observed spectra. A similar approach was used by Andrews (1986) for simultaneous inversion of source spectra, local site amplification effects, and attenuation effects. The CS method uses all the spectra from an event to simultaneously invert for t^*, the near-receiver term (R in equation 1), and a common spectrum for the event.

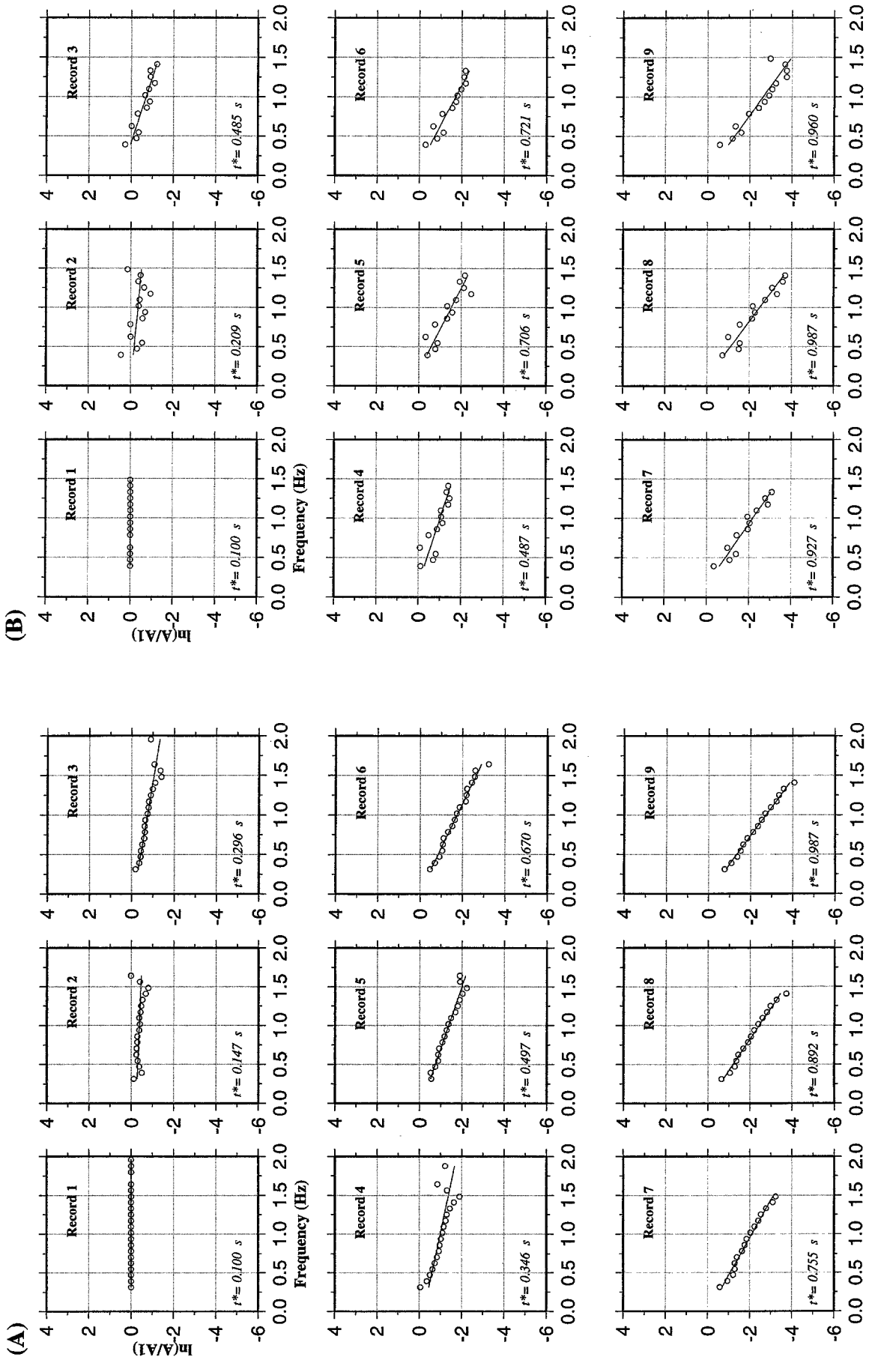
After the seismograms were corrected to a standard response, equations (1) and (2) indicate that the spectrum recorded by the ith station from the event $A_i(f)$ can be expressed as

$$A_i(f) = C(f)R_i(f) \exp(-\pi t_i^* f), \tag{7}$$

where f is frequency, $C(f)$ is the common spectrum for the event, and $R_i(f)$ is the near-receiver effects that are assumed to be frequency and source-location independent. Unlike the approach used by Andrews (1986), the original form of (7) rather than its logarithmic form is used in CS, in order to avoid overweighting near-zero data points, as suggested by Halderman and Davis (1991).

We use a Bayesian approach for the nonlinear inversion (Tarantola and Valette, 1982; Matsu'ura and Hirata, 1982; Jackson and Matsu'ura, 1985) to search for t_i^*, R_i, and $C(f)$. The Bayesian approach uses *a priori* information to choose the starting parameters and constrain the final solution to be within *a priori* bounds. A brief summary of the procedure used in this study is found in the Appendix. Data processing parameters affecting the results include the following:

T, T_0, SN, bp_1, bp_2: same as in the SR method;
a_0: starting estimates for $C(f)$, R_i, and t_i^*. We use the mean spectrum of all the records from a given event as the starting estimates for $C(f)$, R_i values are initially set equal to 1, and the starting estimates for the t_i^* values


(A)
(B)
ln(A/A1)
Frequency (Hz)
Record 1
t*= 0.100 s
Record 2
t*= 0.147 s
Record 3
t*= 0.296 s
Record 4
t*= 0.346 s
Record 5
t*= 0.497 s
Record 6
t*= 0.670 s
Record 7
t*= 0.755 s
Record 8
t*= 0.892 s
Record 9
t*= 0.987 s
Record 1
t*= 0.100 s
Record 2
t*= 0.209 s
Record 3
t*= 0.485 s
Record 4
t*= 0.487 s
Record 5
t*= 0.706 s
Record 6
t*= 0.721 s
Record 7
t*= 0.927 s
Record 8
t*= 0.987 s
Record 9
t*= 0.960 s

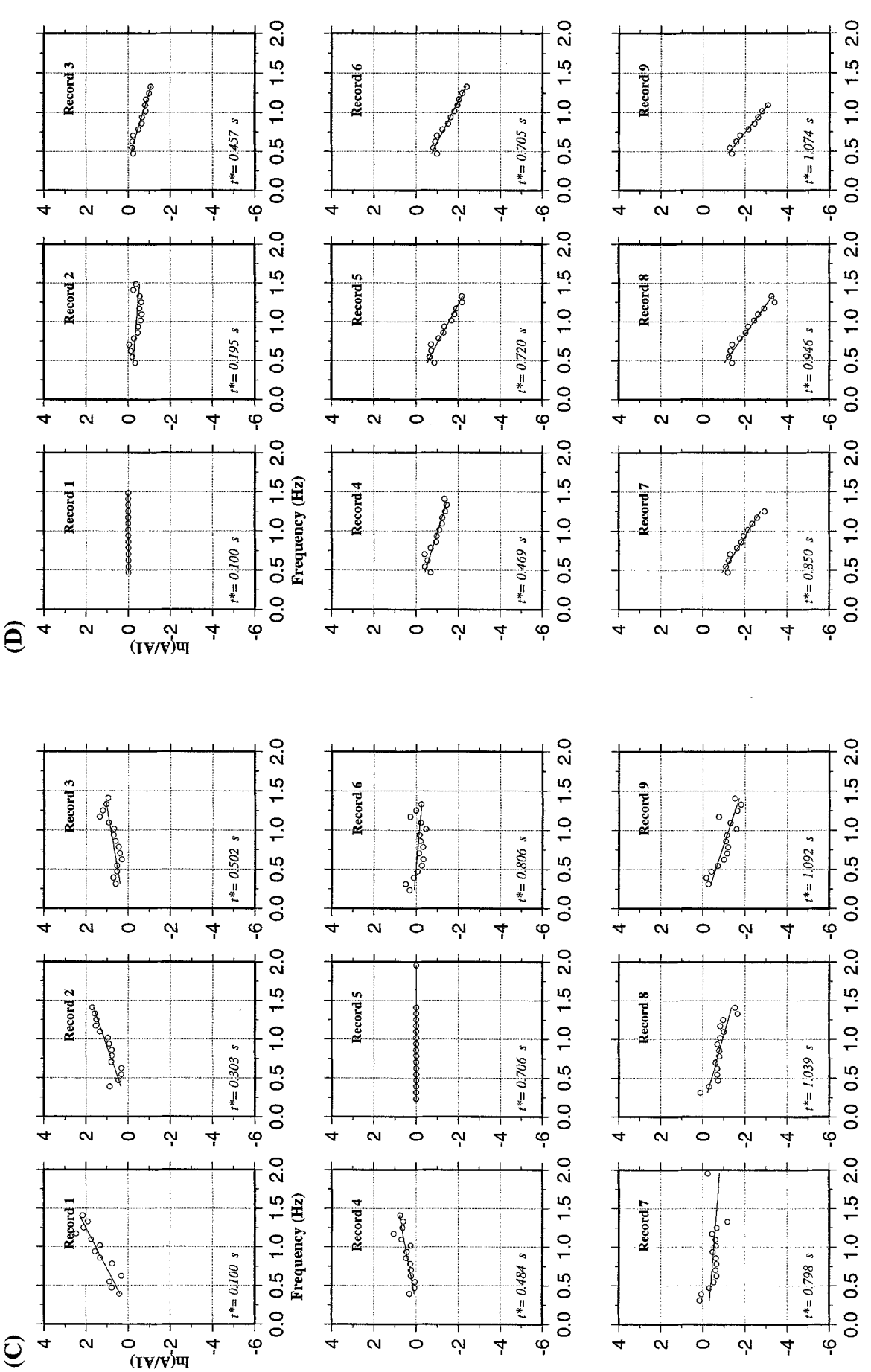


Figure 3. Examples of "observed" (circles) and fitted spectral ratios (line) obtained using the spectral ratio method. The resulting t^* values shown at the lower-left corners of each diagram are relative to the reference record. For easy comparison with the true t^* values, a certain number was added to the resulting t^* values so that the t^* for record 1 is always 0.1 sec. In (a), $z = 0.1$, $F_s = 2$ points, and record 1 is the reference record; in (b), $z = 0.3$, $F_s = 2$ points, and record 1 is the reference record; in (c), $z = 0.3$, $F_s = 2$ points, and record 5 is the reference record; and in (d), $z = 0.3$, $F_s = 4$ points, and record 1 is the reference record.

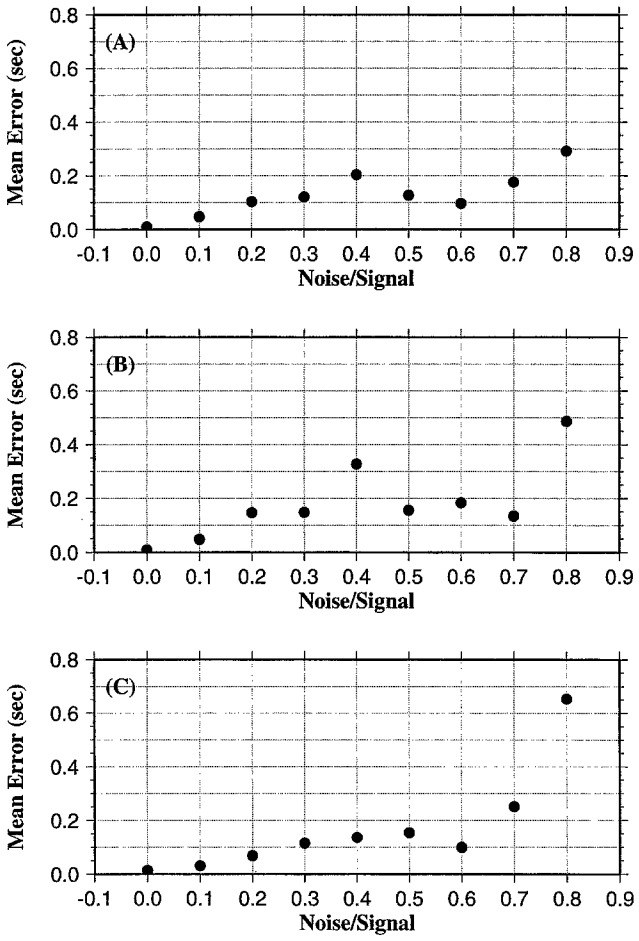


Figure 4. Mean measurement errors over all nine records from spectral ratio method at different noise levels. In (a), F_s = 2 points, and record 1 is the reference record; in (b), F_s = 2 points, and record 5 is the reference record; and in (c), F_s = 4 points, and record 1 is the reference record.

are taken as the values found by the spectral ratio method mentioned above.

σ_0: standard deviations of a_0. Together with a_0, the selection of σ_0 is the most important step in the Bayesian inversion. It reflects the confidence limits for a_0. A parameter with a larger σ_0 is more likely to be modified during the inversion than the ones with smaller σ_0. Because we have large uncertainties on t^* values computed by the spectral ratio method, a large standard deviation (0.5 sec) is used for all nine records. The standard deviations for $C(f)$ are taken as 10, which is about 100% of the peak value of the mean spectrum. For R_i, we use σ_0 = 0.1.

N_{loop}: the number of iterations for the inversion.

Note that F_s, the length of frequency window for smoothing is not used in this method; that is, no smoothing is performed, and no reference record is needed. In this study, we use nine synthetic seismograms, each with 24 frequency points. Therefore, the number of data points in the inversion is 216. The total number of parameters is 42, which includes 24 data points in $C(f)$, 9 R_i's, and 9 t_i^*'s.

Examples of observed and fitted Fourier spectra are shown in Figure 5. For z = 0.1, the match is excellent with a final sum of squared residuals less than 0.5 for all of the records. For z = 0.3, the resulting t^*'s are much closer to the true values than those from the SR method (Figs. 5b, 3b, and 3c).

Measurement errors increase with z, as indicated by the mean measurement errors for different z values shown in Figure 6. All the measurement errors are less than 0.08 sec. Most of the errors are less than ½ of those from the SR method (Fig. 6b).

Discussion

The main reason for the obvious advantage of the CS method over the SR method is that the latter method is strongly affected by the spectrum of the reference record. On the other hand, by finding and simultaneously using a common spectrum, the level of random noise is reduced to about $1/\sqrt{n}$ in the common spectrum, where n is the number of records used.

If the shape of the source spectrum can be roughly estimated by using recordings from near-source stations or other methods, the CS method can be used to fine-tune the source spectrum and to calculate the absolute t^* along the entire path using data from a local or regional seismic network. The starting values for $C(f)$ in equation (7) are now the estimated source spectrum, and their σ values are related to the uncertainties in the initial estimates. A similar approach has been used by Andrews (1986) to determine source parameters. Others (e.g., Hartzell, 1992; Kato *et al.*, 1995) used the idea of Andrews (1986) in the study of local site effects on direct P and S waves, by using all the recordings from all the events in a single inversion.

The CS method assumes that before a seismic wave reaches the uppermost layer of interest, the spectrum of the arriving wave is the same (or very similar) beneath all the stations in a network. If the dimension of the network is small, for example, comparable to the thickness of the crust, the effects of Q heterogeneity in the subcrustal mantle are small, and the observed δt^* values should be small and related to crustal effects. On the other hand, δt^* values observed using a regional network are usually large and are mostly related to mantle effects.

The CS method assumes that both t^* and R are frequency independent and that all the instruments are normalized to a standard response; that is, $I_i(f)$ in equation (1) is the same for all the stations in the array. Occasionally, these assumptions do not reflect the reality. For instance, in some cases, it is found that Q is strongly dependent on frequency for high-frequency body waves (Anderson and Given, 1982). Local site effects can also be frequency dependent, such as reverberations in sedimentary layers. If

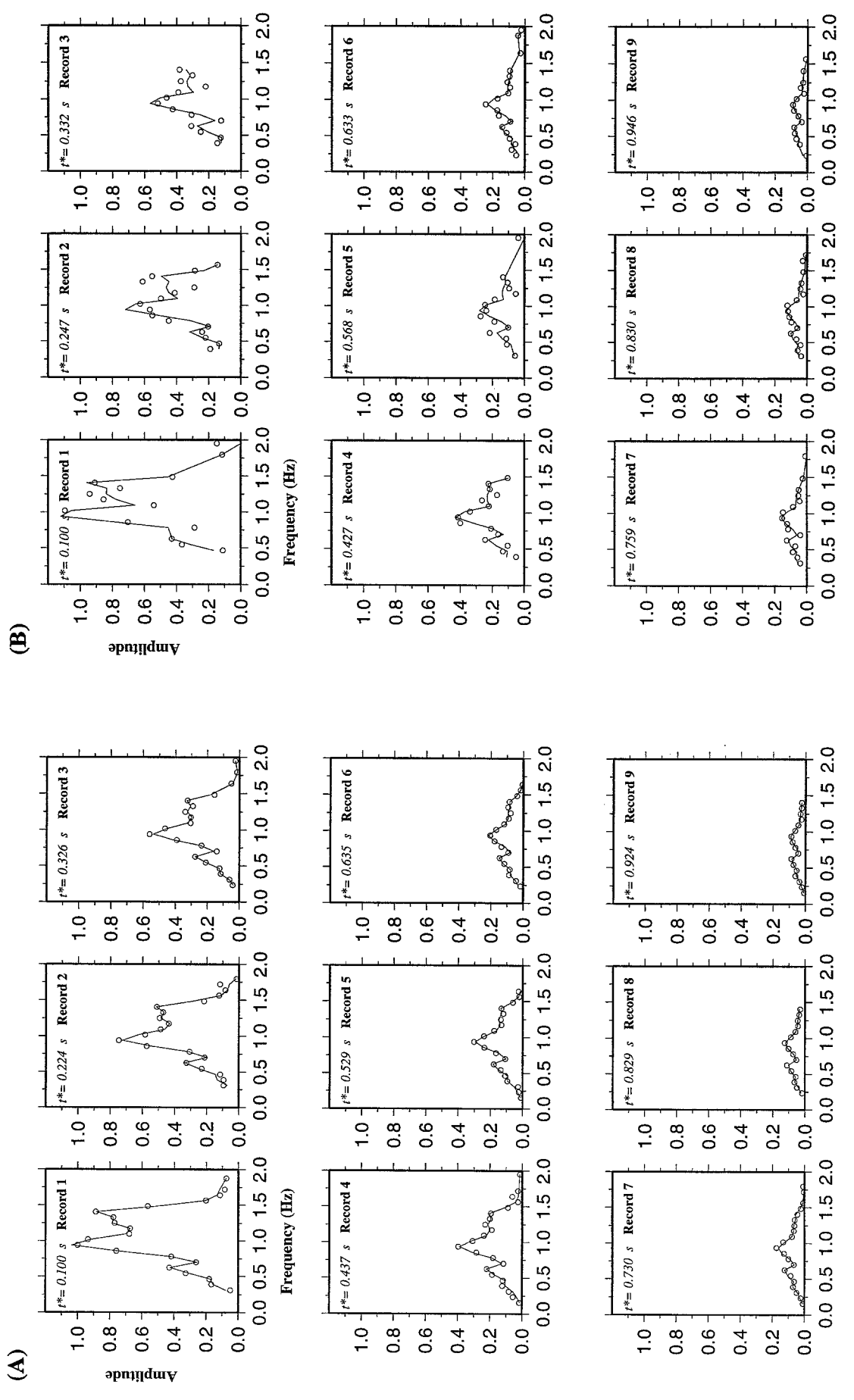


Figure 5. "Observed" (circles) and fitted (line) Fourier spectra using the CS method. A certain number was added to the resulting t^* values shown at the upper-left corners of each diagram, so that the t^* for record 1 is always 0.1 sec. (a) is for $z = 0.1$, and (b) is for $z = 0.3$. Note that only data points with SN $\geqq$ 2.0 are plotted and used in the fitting.

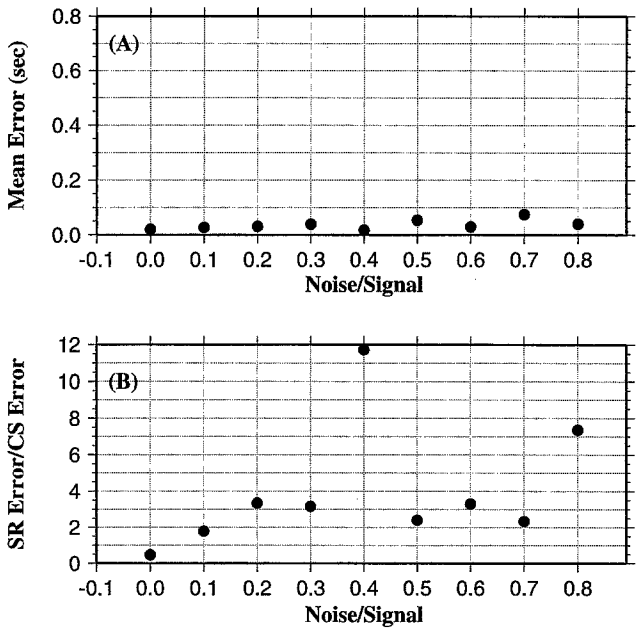


Figure 6. (a) Mean measurement errors over all nine records from the CS method at different noise levels. (b) Ratios of the errors between those shown in Figures 4(a) and 6(a).

some records within a network are strongly affected by such effects, $C(f)$ estimated using all the records is biased. A biased $C(f)$ would lead to errors in the estimation of t^* even for stations with frequency-independent t^*'s. To solve this problem, we propose a multi-step Bayesian inversion procedure. At each step of the inversion, a certain number of iterations are performed. Records with a misfit larger than a cutoff value are rejected from the next step. The underlying assumption for this multi-step procedure is that t^* and R beneath most of the stations are frequency independent and that the instrumental response for most of the stations is accurately corrected. Preliminary results using data from the Baikal seismic array (e.g., Gao, 1995) indicate that these assumptions are acceptable. For a typical event, about 20% of the stations were rejected after the first round of iteration. We found that a station rejected by one event was usually rejected by other events, implying that at least one of the three parameters [R, t^*, and $I(f)$] differs in a way that violates the assumptions of the CS method.

Acknowledgments

I would like to thank Paul Davis for comments on an early version of the manuscript. This study benefited greatly from a course, "Geophysical Data Inversion Theory," conducted by David D. Jackson at UCLA. Reviews by Philip Carpenter and an anonymous reviewer improved the article greatly. Cooperation and encouragement provided by Hong Liu and Annie Gao are gratefully acknowledged. This study was supported by DARPA under Contract F49620-94-1-0161. The revision was done at DTM, Carnegie Institution of Washington while I was a postdoctoral associate supported by NSF under Grant EAR 95-26840.

References

Anderson, D. L. (1967). The inelasticity of the mantle, *Geophys. J. R. Astr. Soc.* **14,** 135–164.

Anderson, D. L. and J. W. Given (1982). Absorption band Q model for the Earth, *J. Geophys. Res.* **87,** 3893–3904.

Andrews, D. J. (1986). Objective determination of source parameters and similarity of earthquakes of different size, *Geophys. Monograph,* Vol. 37, American Geophysical Union, Washington, D.C., 259–267.

Bevington, P. R. (1969). *Data Reduction and Error Analysis for the Physical Sciences,* McGraw-Hill, New York, 237–239.

Der, Z. A. and T. W. McElfresh (1976). The effect of attenuation on the spectra of P waves from nuclear explosions in North America, SDAC-TR-76-7, Teledyne Geotech, Alexandria, Virginia.

Der, Z. A., R. P. Masse, and J. P. Gurski (1975). Regional attenuation of short-period P and S waves in the United States, *Geophys. J. R. Astr. Soc.* **40,** 85–106.

Frankel, A. and W. Leith (1992). Evaluation of topographic effects on P-waves and S-waves of explosions at the Northern Novaya Zemlya test site using 3-D numerical simulations, *Geophys. Res. Lett.* **19,** 1887–1890.

Gao, S. (1995). Seismic evidence for small-scale mantle convection under the Baikal rift zone, Siberia, *Ph.D. Thesis,* University of California, Los Angeles, 221 pp.

Gao, S., H. Liu, P. M. Davis, and L. Knopoff (1996). Localized amplification of seismic waves and correlation with damage due to the Northridge earthquake: evidence for focusing in Santa Monica, *Bull. Seism. Soc. Am.* **86,** S209–S230.

Halderman, T. P. and P. M. Davis (1991). Q_p beneath the Rio Grande and east African rift zones, *J. Geophys. Res.* **96,** 10113–10128.

Hartzell, S. H. (1992). Site response estimation from earthquake data, *Bull. Seism. Soc. Am.* **82,** 2308–2327.

Jackson, D. D. and D. L. Anderson (1970). Physical mechanisms of seismic-wave attenuation, *Rev. Geophys. Space Phys.* **8,** 1–63.

Jackson, D. D. and M. Matsu'ura (1985). A Bayesian approach to nonlinear inversion, *J. Geophys. Res.* **90,** 581–591.

Kato, K., K. Aki, and M. Takemura (1995). Site amplification from coda waves: validation and application to S-wave site response, *Bull. Seism. Soc. Am.* **85,** 467–477.

Knopoff, L. (1964). *Q. Rev. Geophys.* **2,** 625–660.

Matsu'ura, M. and N. Hirata (1982). Generalized least-squares solutions to quasi-linear inverse problem with *a priori* information, *J. Phys. Earth* **30,** 451–468.

Press, W. H., S. A. Teukolsky, W. T. Vettering, and B. P. Flannery (1992). *Numerical Recipes in FORTRAN, the Art of Scientific Computing,* 2nd ed., Cambridge Univ. Press, Cambridge, U.K., 271 pp.

Richards, P. G. and W. Menke (1983). The apparent attenuation of a scattering medium, *Bull. Seism. Soc. Am.* **73,** 1005–1021.

Solomon, S. C. and M. N. Toksöz (1970). Lateral variation of attenuation of p and s waves beneath the United States, *Bull. Seism. Soc. Am.* **60,** 819–838.

Tarantola, A. and B. Valette (1982). Inverse problems: quest for information, *J. Geophys.* **50,** 159–170.

Taylor, S. R., B. P. Bonner, and G. Zandt (1986). Attenuation and scattering of broadband P and S waves across North America, *J. Geophys. Res.* **91,** 7309–7325.

Teng, T. L. (1968). Attenuation of body waves and the Q Structure of the mantle, *J. Geophys. Res.* **73,** 2195–2208.

Vidale, J. E., O. Bonamassa, and H. Houston (1991). Directional site resonances observed from the 1 October 1987 Whittier Narrows, California, earthquake and the 4 October aftershock, *Earthquake Spectra* **7,** 107–125.

Appendix
Nonlinear Bayesian Inversion

The method that we used to search for the optimal parameters is nonlinear Bayesian inversion (Tarantola and

Valette, 1982; Matsu'ura and Hirata, 1982; Jackson and Matsu'ura, 1985). This section summarizes the theory and describes the procedure used in this study.

The relations shown in equation (7) can be generalized as

$$y = f(x, t) + e, \tag{A1}$$

where y is an n-vector of observed data, x is an m-vector of unknown parameters to be found, t is an independent variable, f is an n-vector of known functions, and e is an n-vector of Gaussian errors with zero means and known covariance for the observed data. In the study, nine synthetic records were used, each with 24 data points. Therefore, $n = 216$. The unknown parameters include 24 points in the common spectrum, 9 frequency-independent amplification factors, and 9 t^* values. Therefore, $m = 42$.

The Bayesian approach uses some *a priori* information of the form

$$z = g(x, t) + d, \tag{A2}$$

where z is a vector of *a priori* information, g is a vector of known functions of the m unknown parameters, and d is a vector of Gaussian errors with zero means and known covariance for the *a priori* information.

In this study, z is a 42-vector containing preliminary estimates of the parameters, that is, $g(x, t) \equiv x$. Parameters 1 to 9 are relative t^* values determined from the SR method, parameters 10 to 33 are estimates of the common spectrum obtained from averaging of the normalized raw spectra, and parameters 34 to 42 are *a priori* estimates of the local site amplification factors that are initially assigned to be 1.0. The magnitude of d relative to z reflects the uncertainties in the *a priori* estimate. An *a priori* estimate with a large d implies that it is more likely to be modified during the inversion than the ones with smaller d.

Expansion of (A1) and (A2) in Taylor series about the starting parameters $\{x_{01}, x_{02}, \ldots, x_{0m}\}$ gives

$$y_k = f(x_{01}, x_{02}, \ldots, x_{0m}, t_k) + \frac{\partial f}{\partial x_i}\Big|_{x_i = x_{0i}} (x_i - x_{0_i}) + \delta_1 + e_k, \tag{A3}$$

and

$$z_k = g(x_{01}, x_{02}, \ldots, x_{0m}, t_k) + \frac{\partial g}{\partial x_i}\Big|_{x_i = x_{0i}} (x_i - x_{0i}) + \delta_2 + d_k, \tag{A4}$$

where δ_1 and δ_2 are negligible higher-order terms.

Let

$$y'_k = y_k - f(x_{01}, x_{02}, \ldots, x_{0m}, t_k), \tag{A5}$$

$$x'_k = (x_k - x_{0k}), \tag{A6}$$

$$A'_{ij} = \frac{\partial f}{\partial x_j}\Big|_{t = t_i, x = x_j}, \tag{A7}$$

$$z'_k = z_k - g(x_{01}, x_{02}, \ldots, x_{0m}, t_k), \tag{A8}$$

and

$$B'_{ij} = \frac{\partial g}{\partial x_j}\Big|_{t = t_i, x = x_j}; \tag{A9}$$

by ignoring higher-order terms δ_1 and δ_2, (A3) and (A4) can be written as the following linear forms:

$$Y' = A'X' + e, \tag{A10}$$

$$Z' = B'X' + d. \tag{A11}$$

Because in this study $g \equiv x$, B' becomes a unit matrix.

To standardize, we apply two diagonal matrices, F and G, to (A10) and (A11) so that

$$F^T F = E^{-1}, \tag{A12}$$

and

$$G^T G = D^{-1}, \tag{A13}$$

where E and D are covariance matrices satisfying

$$E = \overline{ee^T}, \tag{A14}$$

and

$$D = \overline{dd^T}. \tag{A15}$$

Two new matrices can be formed by letting

$$Y'' = \begin{pmatrix} FY' \\ GZ' \end{pmatrix}, \tag{A16}$$

$$A'' = \begin{pmatrix} FA' \\ GB' \end{pmatrix}. \tag{A17}$$

The problem can then be expressed in the most general linear form

$$Y'' = A''X' + e', \quad e' \sim N(0, I). \tag{A18}$$

The estimated parameters are calculated using

$$\hat{x} = H''Y'' = H'Y' + K'Z', \tag{A19}$$

where the inverse matrices

$$H'' = (A''^T A'')^{-1} A''^T, \quad \text{(A20)}$$

$$H' = (A'^T E^{-1} A' + B'^T D^{-1} B')^{-1} A'^T E^{-1}, \quad \text{(A21)}$$

$$K' = (A'^T E^{-1} A' + B'^T D^{-1} B')^{-1} B'^T D^{-1}. \quad \text{(A22)}$$

Iterations are performed by replacing starting parameters in equation (A3) with $\hat{x}$ in equation (A19) until a satisfactory match between the calculated and the observed data is obtained. About 150 iterations were used in this study.

Department of Earth and Space Sciences
University of California
Los Angeles, California 90095
E-mail: sgao@dtm.ciw.edu
http://www.ciw.edu/sgao/sgao.html

Manuscript received 30 October 1996.