12WCEE 2000

BLIND DECONVOLUTION OF GROUND SURFACE SEISMIC RECORDINGS FOR SITE RESPONSE IDENTIFICATION

A ZERVA[1], A P PETROPULU[2] And P Y BARD[3]

SUMMARY

A novel blind deconvolution methodology for the identification of local site characteristics based on two seismograms recorded on the free surface of a sediment site is presented. The approach considers that the surface recordings are the result of the convolution of the "input motion at depth" with the site response functions ("channels"), that represent the characteristics of the transmission path of the waves from the input location to each recording station. The input motion at depth is considered to be the common component in the seismograms. The channel characteristics are considered to be the part in the seismograms that is noncommon, since the travel path of the waves from the input motion location at depth to each recording station is different, due to spatially variable site effects. By means of blind deconvolution, the algorithm eliminates what is common in the seismograms, namely the input motion at depth, and retains what is different, namely the site response functions from the input location to each recording station. It estimates the site response in both frequency and time domains, and identifies the duration of the site response functions. The methodology is applied herein to synthetic data at realistic sites for performance validation.

INTRODUCTION

It has been shown, through observations and theory, that local site conditions amplify seismic ground motions significantly as compared to those recorded on bedrock. This amplification increases the potential of damage and collapse of engineering structures. The evaluation of site characteristics from analyses of recorded data has attracted recently significant attention; Field and Jacob (1995) and Safak (1997) presented extensive reviews on the subject.

Generally, the available methods for establishing the site specific response using recorded data can be classified as "reference" and "non-reference" site techniques. The most common reference site techniques are based on the evaluation of spectral ratios, that require recorded data at two locations: the first recording (input motion), ideally, at the base of the sediment site and the second (output motion) at its free surface. However, because the true input motion

[1] *Dept of Civil and Architectural Eng, Drexel University, Philadelphia, USA. E-Mail: zervaa@drexel.edu*
[2] *Dept of Electrical and Computer Eng, Drexel University, Philadelphia, USA.*
[3] *Observatoire de Grenoble, Joseph Fourier University, Grenoble Cedex, France.*

cannot be recorded, it is approximated by downhole recordings, which may be contaminated by free surface effects, or by recordings at a nearby rock outcrop, for which it is assumed that the source effects, azimuth and incident angles of waves are the same as those at the base of the site (Field and Jacob, 1995). The non-reference site techniques involve a generalized inversion scheme that estimates source, path and site effects (Boatwright, Fletcher and Fumal, 1991), but require prior parameterization of the source and path characteristics. Alternatively, the receiver function technique, i.e., horizontal-to-vertical spectral ratios of microtremors (Nakamura, 1989) or of the S-wave signal (Lermo and Chavez-Garcia, 1993), is a very simple technique that appears to reproduce the overall frequency content of the site, but its advantages, limitations and justification are still under investigation (Lachet and Bard, 1994).

This paper employs a recently proposed blind deconvolution methodology (Pozidis and Petropulu, 1997; Zerva, Petropulu and Bard, 1999) for the reconstruction of the seismic input motion at depth and the estimation of the site characteristics through response functions (channels) between the seismic input location and the two stations. The required data are seismograms recorded at two different stations on the free surface of the sediment site. In this way, the approach eliminates the need for either a recording at a nearby rock site or a downhole recording and does not require any prior parameterization of source and path effects. The methodology, which is briefly described in the following section, is applied herein to synthetic data evaluated at realistic sites for performance validation.

2. BLIND DECONVOLUTION METHODOLOGY

Let $x_1(n)$ and $x_2(n)$ be two seismograms recorded at two free surface stations; it is considered herein, as in most linear site response analyses, that:

$$x_i(n) = h_i(n) * s(n) + \eta_i(n), \qquad i = 1, 2 \tag{1}$$

i.e., the surface recordings $x_1(n)$ and $x_2(n)$ are the convolution of the motion at depth ("input"), $s(n)$, with the site response functions ("channels"), $h_1(n)$ and $h_2(n)$, representing the transmission of the waves from the input location to the each recording station; n in Eq. 1 indicates discrete time, i.e., $n = t/\Delta t$, t being time and Δt the time step; "$*$" indicates convolution, and $\eta_1(n)$ and $\eta_2(n)$ indicate noise at the two locations.

The input motion $s(n)$, which is considered to be the common component in the seismograms (same input in a statistical sense), is described by a stationary, zero-mean random process:

$$s(n) = e(n) * r(n) \tag{2}$$

with $e(n)$ indicating a zero-mean white process, and $r(n)$ the color of the process. It is noted that the terms "input motion" and its "location at depth" are vague: Because the approach identifies these parameters blindly, the actual input motion and its location are not known *a priori*. In the ideal case, the input motion would be the incident motion at the bedrock-surface layer(s) interface, and the location would be that interface. However, depending on parameters, as, e.g., the recording stations' separation distance, the "location" of the input motion could be at an intermediate depth within the layers, and the input motion would then incorporate the effects of the lower layers.

The channels, $h_1(n)$ and $h_2(n)$ in Eq. 1, are considered to be the part in the seismograms that is noncommon, since the travel path of the waves from the input motion location to each recording station is different, due to spatially variable site effects. The blind deconvolution algorithm eliminates what is common in the seismograms, namely the input motion at depth, $s(n)$, and retains what is different, namely the site response functions from the input location to each recording station. The approach considers that the channels are nonminimum phase. The analysis based on nonminimum phase channels is more general and includes the minimum

phase channels as a special case. For both cases the amplitude spectra of the sequences will be identical, but their phase spectra will be different. Waiving the minimum phase restriction on the channels allows the evaluation of the true time domain characteristics of the site response. The channels are then expressed in terms of their minimum, $i_i(n)$, and maximum, $o_i(n)$, phase components as:

$$h_1(n) = i_1(n) * o_1(n), \;\; h_2(n) = i_2(n) * o_2(n) \tag{3}$$

Finally, by means of deconvolution of each seismogram with the corresponding estimated channel, the input motion can also be reconstructed.

The blind deconvolution methodology (Pozidis and Petropulu, 1997) proceeds then as follows: From Eqs. 1 and 2, the recorded sequences are rewritten as

$$x_i(n) = e(n) * g_i(n) + \eta_i(n), \qquad i = 1, 2 \tag{4}$$

where $g_i(n)$ is the convolution of the color of the input, $r(n)$, with the respective channel, $h_i(n)$, $i = 1, 2$. Let $\tilde{g}_i(n)$ denote the minimum-phase equivalent (Oppenheim and Schafer, 1989) of $g_i(n)$, and $\tilde{G}_i(\omega)$, ω indicating discrete frequency, its Fourier transform. The minimum phase equivalent is obtained from the power cepstrum, $C_i(\omega)$, $i = 1, 2$, of the recorded data:

$$\tilde{g}_i(n) = F^{-1}\{exp[C_i(\omega)]\}; \;\; c_i(n) = F^{-1}\{ln[S_{x_i}(\omega)]\}u(n) \tag{5}$$

where, $S_{x_i}(\omega)$ is the power spectrum of $x_i(n)$, $c_i(n)$ is the Fourier transform of $C_i(\omega)$, $u(n)$ equals one for $n > 0$ and zero otherwise, and F^{-1} indicates inverse Fourier transform. Let

$$S_{min}(\omega) = S_{x_1x_2}(\omega)\tilde{G}_1(\omega)\tilde{G}_2^*(\omega); \;\; S_{max}(\omega) = S_{x_1x_2}(\omega)\tilde{G}_1^*(\omega)\tilde{G}_2(\omega) \tag{6}$$

where, $S_{x_1x_2}(\omega)$ is the cross-spectrum of $x_1(n)$ and $x_2(n)$, and superscript * indicates conjugation. Based on the phase properties of $\tilde{g}_1(n)$ and $\tilde{g}_2(n)$ in relation to those of the convolutional components of $x_1(n)$ and $x_2(n)$ (Eq. 4), it can be shown (Pozidis and Petropulu, 1997) that:

$$arg\{S_{i_1i_2}(\omega)\} = \frac{1}{2}arg\{S_{min}\} + \frac{1}{2}k\omega; \;\; arg\{S_{o_1o_2}(\omega)\} = \frac{1}{2}arg\{S_{max}\} + \frac{1}{2}k\omega \tag{7}$$

where k is an integer, $S_{i_1i_2}(\omega)$ is the cross-spectrum of the minimum phase parts of $h_1(n)$ and $h_2(n)$, and $S_{o_1o_2}(\omega)$ is the cross-spectrum of the maximum phase parts of $h_1(n)$ and $h_2(n)$; the contribution of the minimum, $i_i(n)$, and maximum, $o_i(n)$, phase parts to the respective channels, $h_i(n)$, has been described in Eq. 3. Let

$$h_{min}(n) \triangleq i_1(n) * i_2^*(-n); \;\; h_{max}(n) \triangleq o_1(n) * o_2^*(-n) \tag{8}$$

be the time equivalents of $S_{i_1i_2}(\omega)$ and $S_{o_1o_2}(\omega)$, respectively (Eqs. 7). It can be shown (Pozidis and Petropulu, 1997), that the sequences $h_{min}(n)$ and $h_{max}(n)$ can be reconstructed from the phase of $S_{i_1i_2}(\omega)$ and $S_{o_1o_2}(\omega)$, respectively, within a scalar.

When the phase and the duration (length) of the sequence are known exactly, the sequence can be obtained as the least-squares solution of a system of equations, which involves known phase samples, i.e.,

$$\sum_{n=1}^{N-1} x(n)sin(\omega n - \theta_x(\omega)) = sin(\theta_x(\omega)), \{\omega = \frac{2\pi}{L}k, \;\; k = 0, .., L-1\} \tag{9}$$

where $x(n)$ plays the role of $h_{min}(n)$ or $h_{max}(n)$, $\theta_x(\omega)$, N are the corresponding Fourier phase and length, and it was assumed that $x(0) = 1$. However, from Eqs. 7, the phases of $S_{i_1i_2}(\omega)$ and $S_{o_1o_2}(\omega)$ can be computed up to a linear phase, and the length of the sequences is not known. Therefore, in order to identify $h_{min}(n)$ and $h_{max}(n)$, the reconstruction from phase methodology has to be applied inside a double loop as explained in the sequel. Let

L_{min} be the true length of $h_{min}(n)$, which is unknown, and let l be a guess for it; then: For $l = 2, 3, ...,$ and within each l loop, for $m = -l : 0.5 : l$ (i.e., m increases between $-l$ and l with step 0.5): Rewrite Eq. (9) based on the phase $\theta_m(\omega) = arg\{S_{min}(\omega)\}/2 - m\omega$, and length l. Let $\mathbf{x}_{m,l}$ be the corresponding least-squared solution, and $\theta_{x_{m,l}}(\omega)$ be its Fourier phase. If $e_{min}(\omega) \triangleq arg\{S_{min}(\omega)\}/2 - \theta_{x_{m,l}}(\omega)$ is a straight line in ω, then $\mathbf{x}_{m,l}$ is the right solution, i.e., $x_{m,l}(n) = h_{min}(n)$, and the iteration stops here. If not, then the loop in m proceeds for the next value of m. To determine whether $e_{min}(\omega)$ is a straight line, we perform a least-squares (LS) fit of $e_{min}(\omega)$ to the equation of a line. Let $E_{min}(l, m)$ be the LS error of the fit corresponding to length l and phase $\theta_m(\omega)$, and let $E(l) = \min_m\{E_{min}(l, m)\}$. For each value of l, the reconstructed sequence corresponding to the location of the minimum in $E_{min}(l, m)$ is a potential solution. However, for $l < L_{min}$ no solution exists. For $l \geq L_{min}$, the LS error exhibits a profound minimum that stays almost constant as l increases, and which is several orders of magnitude smaller than the minima corresponding to $l < L_{min}$. We can monitor $E(l)$ as l increases, and when it stabilizes to a low value, i.e., $E(l_0) \approx E(l_0 + 1) \approx E(l_0 + 2)...$ choose as solution the one corresponding to length l_0. It has been shown by Pozidis and Petropulu (1997), that the smallest length solution, indicated by an associated low least-squares error, is the correct solution.

The reconstructed sequence $h_{min}(n)$ (Eqs. 8) is then decomposed into the minimum phase parts of the channels, $i_1(n)$ and $i_2(n)$, and $h_{max}(n)$ (Eqs. 8) into the maximum phase parts of the channels, $o_1(n)$ and $o_2(n)$. This decomposition is achieved through polynomial rooting, so that the minimum and maximum phase zeros are determined.

Finally, from the information on their minimum and maximum phase parts, the channels are computed either in the time domain from Eq. 3, or in the frequency domain as

$$H_1(\omega) = I_1(\omega)O_1(\omega), \quad H_2(\omega) = I_2(\omega)O_2(\omega) \tag{10}$$

where $I_i(\omega)$ and $O_i(\omega)$ are the Fourier transforms of $i_i(n)$ and $o_i(n)$, $i = 1, 2$, respectively. It is noted that, because the true length of the sequences $h_{min}(n)$ and $h_{max}(n)$ is identified in the process, the time domain characteristics of the channels (Eq. 3) reflect their true duration.

3. APPLICATION TO SYNTHETIC SEISMIC DATA

The blind deconvolution methodology described in the previous section is applied herein to synthetic seismic data for performance validation. Two sites with similar properties (Tables 1 and 2) are used in the evaluation of the synthetic seismograms. The sites consist of horizontal layers overlaying the bedrock, which has the same properties in both cases. Site 1 (Table 1) consists of five horizontal layers over the bedrock; the characteristics of this site (layers and bedrock) are essentially those reported for Ashigara Valley, station KS2 (Sawada, 1992). Site 2 consists of three layers overlaying the bedrock; the total depth of the layers is the same as that of Site 1, and the layer properties of Site 2 are rough average values of those of Site 1. Vertical incidence is considered for the transmission of the waves from the bedrock to the free surface of each site. The incident motion at the bedrock-layers interface is a segment (110 sample points) of the time history recorded during the strong motion shear wave window at station I03 of the SMART-1 array in Lotung, Taiwan, during the earthquake of January 29, 1981 ($M_L = 6.3$). A small amount of Gaussian noise, uncorrelated at the two stations, is also added to the time histories. Δt in this example is 0.02sec. The synthetic time and frequency domain results at the free surface of the two sites are presented in Fig. 1, and are considered to be the "two recorded seismograms on the free surface of a sediment site". (It is noted that, in the figures, all quantities have been normalized: since the algorithm reconstructs the site response within a constant, amplitudes in both time and frequency domains have been normalized with respect to their maximum value.)

Table 1: Properties of Site 1 (Ashigara Valley, station KS2)

Layer	Depth m	Density g/cm^3	P-wave Velocity m/sec	S-Wave Velocity m/sec
1	0.0 – 4.0	1.40	1500.0	70.0
2	4.0 – 11.0	1.50	1500.0	150.0
3	11.0 – 78.0	1.70	1800.0	400.0
4	78.0 – 95.0	1.90	2000.0	700.0
5	95.0 – 280.0	2.00	2200.0	800.0
6	280.0–	2.50	3000.0	1500.0

Table 2: Properties of Site 2

Layer	Depth m	Density g/cm^3	P-wave Velocity m/sec	S-Wave Velocity m/sec
1	0.0 – 20.0	1.40	1500.0	150.0
2	20.0 – 120.0	1.70	1800.0	400.0
3	120.0 – 280.0	2.00	2200.0	800.0
4	280.0–	2.50	3000.0	1500.0

From the information in Fig. 1, the algorithm is asked to blindly evaluate the characteristics of the sites in both time and frequency domains. Figure 2 (left column) presents the minimum and maximum phase sequences of the "recorded data", that contain, correspondingly, the minimum and maximum phase components of both channels, as well as those of the incident motion at the bedrock-layers interface; these sequences were constructed through polynomial rooting of the two time histories. Additionally, Fig. 2 (right column) presents the reconstructed minimum and maximum phase sequences of the channels (site response), $h_{min}(n)$ and $h_{max}(n)$, that were identified from the smallest length solution associated with low least-squares error in the reconstruction from phase methodology. In the process of reconstruction of the minimum and maximum phase channel sequences, $h_{min}(n)$ and $h_{max}(n)$, the components of the incident motion have been eliminated. $h_{min}(n)$ and $h_{max}(n)$ of Fig. 2 (right column) are then used for the identification of the minimum and maximum phase components, $i_i(n)$ and $o_i(n)$, $i = 1, 2$, of the channels (Eqs. 8). Figure 3 presents the comparison of the actual impulse response function of Sites 1 and 2 and their Fourier amplitudes with the channel characteristics reconstructed by means of blind deconvolution. It is noted that the blindly estimated results are compared with the impulse response of each site, because, from the set-up of this example, it follows that the "input motion" (common component) is the incident motion at the bedrock-layers interface, and its location is the interface. The agreement of the actual and reconstructed site characteristics in Fig. 3 is almost perfect both in the time and frequency domains. In particular, it is noted that the duration of the actual impulse response functions is fully matched by those of the reconstructed ones. The frequency content of the site response characteristics is also fully captured by those of the blindly identified ones. Only small differences are observed: the blindly estimated site responses are slightly richer in frequency content than the actual site characteristics in the higher frequency range. Figure 4 presents the frequency domain characteristics of the incident motion at the bedrock-layer interface (common component) identified at the two sites. For the evaluation of the common component, the recorded data at each site were deconvolved with the corresponding site characteristics, and, hence, an incident motion was identified at each site. In Fig. 4, the "input motions" at the bedrock-layer interfaces of the two sites fully coincide, as should be expected from the comparison of the estimated and actual site response characteristics (Fig. 3) and the definition of the common component.

4. CONCLUSIONS

A novel blind deconvolution approach for the identification of site response characteristics based on two seismograms recorded on the free surface of the sediment site is presented. The approach considers that the seismic surface motions result from the convolution of a common input excitation at depth with the site response functions, that represent the characteristics of the travel path of the waves from the input motion location to each recording station. It is based on the reconstruction of signals from phase information only, and allows the site response to be non-minimum phase. Unique features of the approach include: (i) it does not require seismic records at depth nor at a near-by rock site, as is commonly the case in site response analyses, and eliminates the need of any prior parameterization of source or site characteristics; (ii) it provides a two-dimensional picture of the site, since it identifies channel characteristics from the "input motion location" to each recording station; current approaches estimate a one-dimensional site response; and (iii) it fully identifies the time domain characteristics (response functions) of the site, including their actual duration. The application of the blind deconvolution methodology to synthetic data using realistic site characteristics and incident motions at the bedrock-layer interface resulted in an essentially perfect match of the actual and the blindly estimated site response characteristics. As any new approach, the proposed algorithm requires further testing before it is fully utilized in actual recorded seismograms; currently, it is being applied to additional synthetic data. The present results strongly indicate that it will yield a promising new tool in seismic site response identification from recorded data.

6. ACKNOWLEDGMENTS

This study was conducted while the first author was stationed at the California Institute of Technology and supported by USA NSF grants CMS-9725567 and POWRE-CMS-9870509; she would like to acknowledge the NSF support, Caltech's hospitality, and many helpful discussions with Professor Hiroo Kanamori of Caltech.

7. REFERENCES

Boatwright, J., Fletcher, J.B. and Fumal, T.E. (1991), "A General Inversion Scheme for Source, Site and Propagation Characteristics Using Multiply Recorded Sets of Moderate-Seized Earthquakes", *Bulletin of the Seismological Society of America* 81, pp1574-1782.

Field, E.H. and Jacob, K.H. (1995), "A Comparison and Test of Various Site Response Estimation Techniques, Including Three that are not Reference-Site Dependent", *Bulletin of the Seismological Society of America* 85, pp1127-1143.

Lachet, C. and Bard, P.-Y. (1994), "Numerical and Theoretical Investigations on the Possibilities and Limitations of the Nakamura's Technique", *Physics of the Earth* 42, pp377-397.

Lermo, J. and Chavez-Garcia, F.J. (1993), "Site Effect Evaluation Using Spectral Ratios with only One Station", *Bulletin of the Seismological Society of America* 83, pp1574-1594.

Nakamura, Y. (1989), "A Method for Dynamic Characteristics Estimation of Subsurface Using Microtremor on the Ground Surface", *QR of Railway Technology Research Institute, Japan*, 30.

Oppenheim, A.V. and Schafer, R.W. (1989), *Discrete-Time Signal Processing*, Prentice Hall.

Pozidis, H. and Petropulu, A.P. (1997), "Cross-Spectrum Based Blind Channel Identification", *IEEE Transactions on Signal Processing*, 45, pp2977-2993.

Safak, E. (1997), "Models and Methods to Characterize Site Amplification from a Pair of Records", *Earthquake Spectra* 13, pp97-129.

Sawada, Y. (1992), (The Subcommittee for the Geotechnical Survey of the Ashigara Valley Blind Prediction Test) "Geotechnical Data", *Proceedings of the International Symposium on the Effects of Surface Geology on Seismic Motion.*

Zerva, A., Petropulu, A.P. and Bard, P.-Y. (1999), "Blind Deconvolution Methodology for Site-Response Evaluation Exclusively from Ground Surface Seismic Recordings", *Soil Dynamics and Earthquake Engineering*, 18, pp47-57.

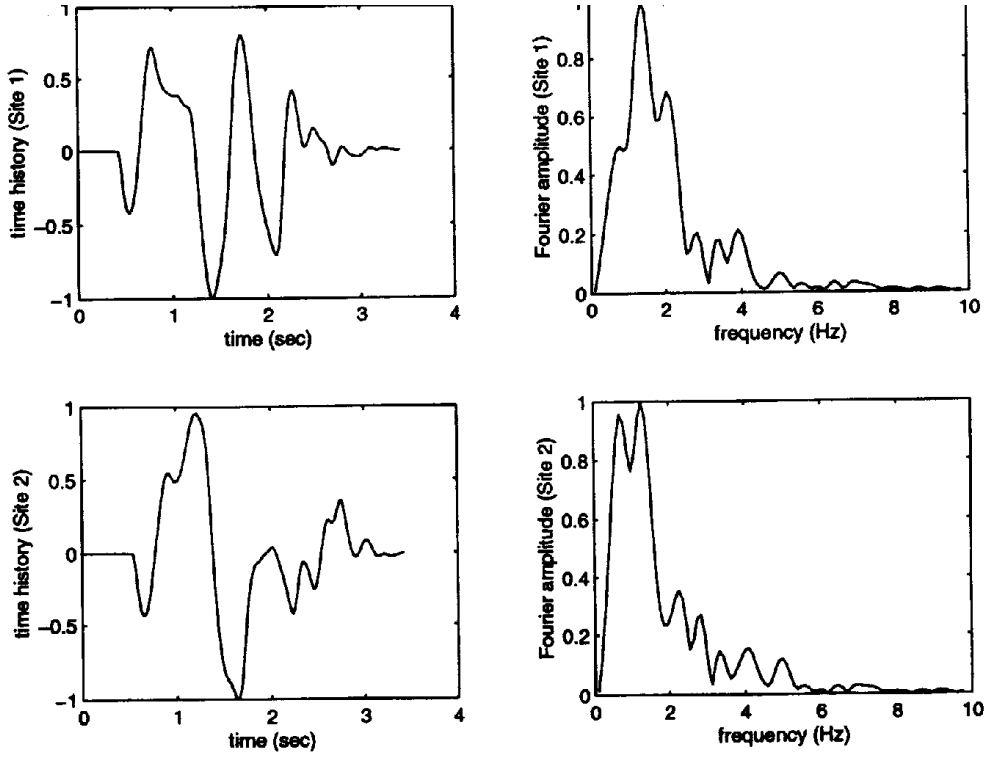

Figure 1: Synthetic motions on the free surface of Sites 1 and 2 in time and frequency domains

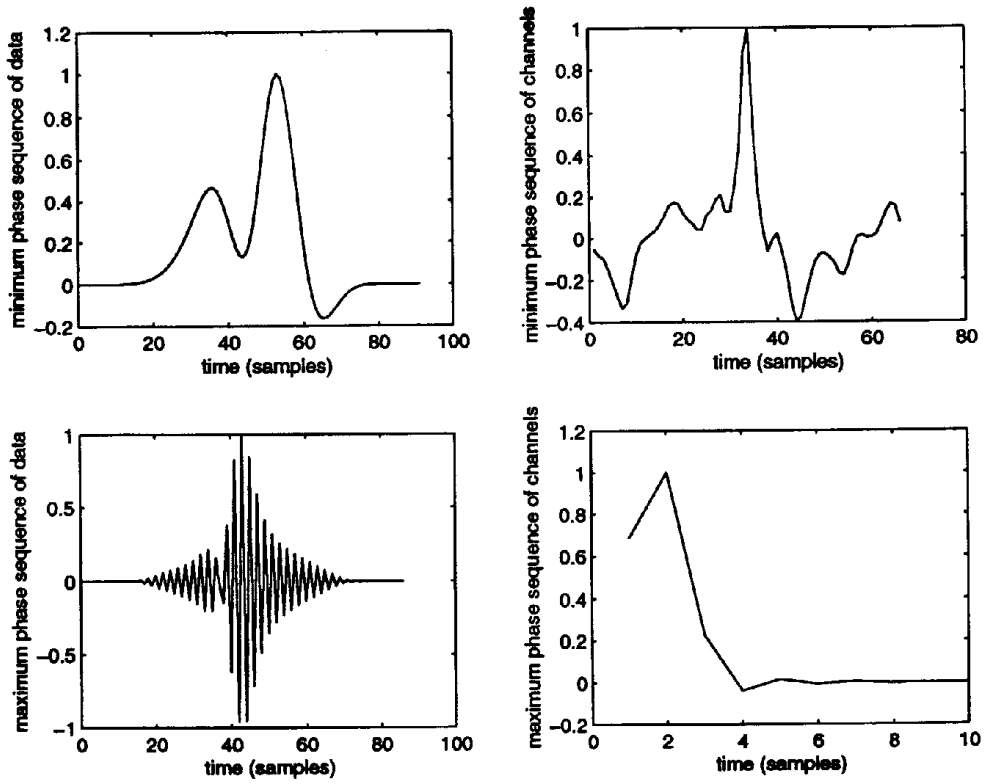

Figure 2: Minimum and maximum phase sequences of the recorded data, and minimum and maximum phase sequences of the channels, $h_{min}(n)$ and $h_{max}(n)$, identified from the reconstruction from phase methodology

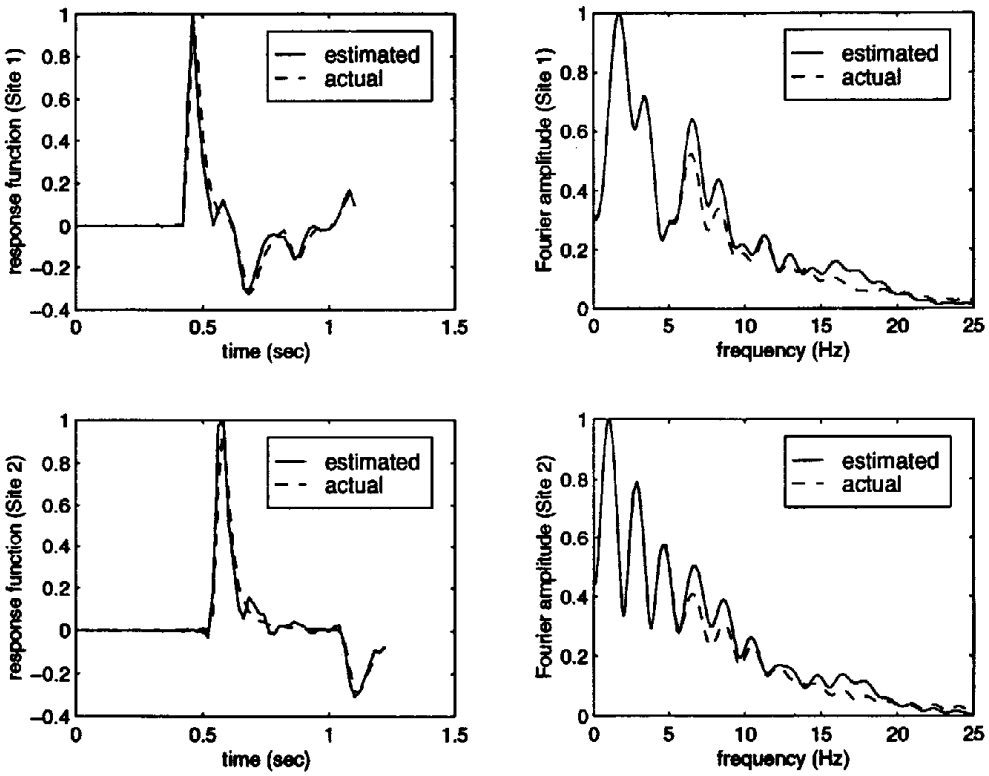

Figure 3: Comparison of actual and blindly estimated response of Sites 1 and 2 in time and frequency domains

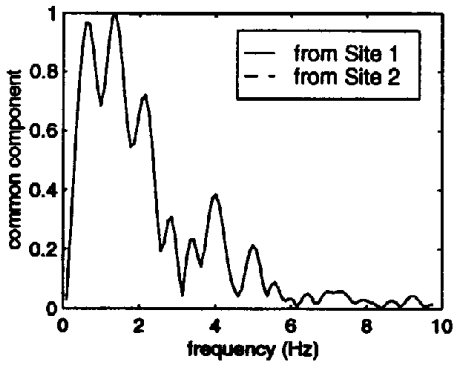

Figure 4: Reconstructed incident motion at the bedrock-layer interface of each site (common component)