Obtaining Spectrum Matching Time Series Using a Reweighted Volterra Series Algorithm (RVSA)

by N. A. Alexander, A. A. Chanerley, A. J. Crewe, and S. Bhattacharya

Abstract In this paper, we introduce a novel algorithm for morphing any accelerogram into a spectrum matching one. First, the seed time series is re-expressed as a discrete Volterra series. The first-order Volterra kernel is estimated by a multilevel wavelet decomposition using the stationary wavelet transform. Second, the higher-order Volterra kernels are estimated using a complete multinomial mixing of the first-order kernel functions. Finally, the weighting of every term in this Volterra series is optimally adapted using a Levenberg–Marquardt algorithm such that the modified time series matches any target response spectrum. Comparisons are made using the SeismoMatch algorithm, and this reweighted Volterra series algorithm is demonstrated to be considerably more robust, matching the target spectrum more faithfully. This is achieved while qualitatively maintaining the original signal’s nonstationary statistics, such as general envelope, time location of large pulses, and variation of frequency content with time.

Introduction

Bounding the uncertainty in predicted ground-motion time series is a complex question that, for most design codes around the world, has concentrated on satisfying a response spectrum of some kind. Code-based response spectra (Biot, 1943; Housner, 1959; Mohraz, 1976) were historically rooted in a simplification of the problem of determining dynamic structural responses to ground excitation. From a structural engineering perspective, the attractiveness of employing a response spectrum was that it was subsequently possible to avoid all time-history analyses. In a sense, the response spectrum tries to characterize the unknown multivariate statistics of seismic ground-motion time series in its own way when convolved through a structural system. Response spectra are smoothed and filtered estimates of the power content of an accelerogram and thus have a practical utility for the engineer and seismologist.

More recently, however, the feasibility and economic benefit of nonlinear time-history analyses of structural/geotechnical systems has become more widespread in design practice. A performance-based design philosophy has led to much more interest in levels of damage at a range of limit states. Nonlinear time-history analyses are required to assess compliance to these various performance-limit states (Building Seismic Safety Council, 1997). This is particularly the case for large infrastructure projects, yet time-history analysis requires ground-motion time series rather than a response spectrum. The problem is that we historically have captured estimates of the unknown multivariate statistics of seismic ground motion in a response spectrum, but a response spectrum does not mathematically imply a unique time series.

Consequently we are presented with the inverse problem of conjecturing a ground-motion time series that, when convolved through a single-degree-of-freedom system, results in a given target response spectrum. If one computes the response spectra for authentic recorded accelerograms, it is apparent that these accelerograms do not individually match a typical code-based spectrum. This is because the code-based spectra are themselves some highly smoothed, mean, curve fit of many individual accelerogram spectra (Japanese Society of Civil Engineers, 1997).

Thus, many researchers (e.g., Ghafory-Ashtiany et al., 2012; Katsanos and Sextos, 2013) in effect advocate a stratified random sampling of recorded ground-motion time series (see Cochran, 1977). Here, a set of real accelerograms are selected so as to satisfy certain geophysical (e.g., event magnitude, epicentral distance, fault mechanism) and structural (e.g., period, ductility demand) criteria. In addition, we propose that this set should have a mean spectrum that is comparable to the target spectra. This approach is implemented on the Pacific Earthquake Engineering Research Center Next Generation Attenuation (PEER NGA) strong-motion database website (PEER, 2010). A set of time series, selected in this way, is then used in time-history analyses of the structural/geotechnical system under consideration. However, the selection of a stratified random sample of records that has a mean spectrum comparable to a target spectrum is difficult to achieve in practice because of the spatiotemporal data sparseness of ground-motion time series. Essentially, we do not have a large enough population of recorded ground-motion time series for all possible geographical locations and all geophysical/structural configurations. Thus, at the very least,
the available records will have to be scaled in amplitude so the mean response spectrum of the stratified sample approximately matches the target spectrum. This amplitude scaling, as Luco and Bazzurro (2007) points out, is not always a completely unbiased modification for the case of nonlinear structural analyses. The hypothesis here is that this stratified and scaled random sample set of records contains a reasonably accurate representation of the population of all significant earthquakes that could occur at a site within a given time frame.

Nevertheless, data sparseness is still a problem in some locations. In most developing countries that are prone to seismic hazard, only a very few recorded motions are available; see, for example, the case for India described in Gobindaraju and Bhattacharya (2012). Thus, there is a need for another approach.

As an alternative, many researchers (Ostadan et al., 1996; Bazzurro and Luco, 2006; Hancock et al., 2006; Cacciola, 2010) advocate the generation of artificial time series that match a given spectrum. This approach is easy to criticize (Priestley, 2003; Al Atik and Abrahamson, 2010), as it is considered overly conservative because these artificial time series have a flat and smooth broadband response spectra that is unlike most real earthquakes. Nevertheless, the perceived attractiveness of spectrally matched records is that fewer accelerograms need to be used in nonlinear time-history analyses. It is worth emphasizing the repetitive and evolutionary nature of design analyses, given the input of various key stakeholders (e.g., client, engineer, architect, and local public planning authorities). Thus, the benefit of reducing run times for computational analyses is not to be underestimated. As an approach, it can be viewed as less sensitive to sample selection, as all matched records should approximate the population statistics of site-specific earthquakes in terms of their spectra. Design codes often require, or at the very least encourage, the use of ground-motion modification. This is predominantly the case when safety-critical facilities are being designed and tested (see Telcordia Technologies, 2002; IEEE Power Engineering Society, 2004, 2005; Takhirov et al., 2005; Crewe, 2012).

Rizzo et al. (1975) discussed a frequency domain technique that alters the Fourier amplitude (but not phase) spectrum of the original seed ground motion so as to match a target response spectrum. This approach suffers because it changes the nonstationary nature of the records. Lilhanand and Tseng (1988) pointed this out and introduced a perturbation technique based on a convolution integral that modifies the original signal into a spectrum-matching one. The added perturbation was localized in time so that it attempted to maintain similar nonstationary statistics to the original. Abrahamson (1992) developed the approach in Lilhanand and Tseng (1988) in producing Rspmatch software. Mukherjee and Gupta (2002) also suggested a wavelet-based modification to the original time series. Rspmatch was further developed by Hancock et al. (2006) in Rspmatch2005. This algorithm is embedded in SeismoSoft Ltd (2012). An improved version of Rspmatch (2010) is described in Al Atik and Abrahamson (2010) that seeks to address some of the problems of nonconvergence of the previous algorithm, although it is still based on the similar perturbation approach of Lilhanand and Tseng (1988).

Even though the term “wavelets” is used in SeismoSoft Ltd (2012) and Hancock et al. (2006), these approaches do not use the wavelet transform commonly employed in the signal processing literature, and they do not make use of conventional mother wavelets. They have not been designed to be optimal filters, and they do not guarantee to only change the phase linearly (see Burrus et al., 1997). In addition, the solution strategy in Lilhanand and Tseng (1988) for determining the optimal adjustment was, in effect, a simple iteration that often suffers from lack of convergence to a reasonable match. Far more robust methods exist in the extensive optimization literature. Having said this, SeismoSoft Ltd (2012) is a valuable resource and allows for a simultaneous match of response spectra across multiple damping values. One of its strengths is that it mostly conserves significant features of the original record while trying to match the target response spectrum.

Aims

In this paper, we seek a novel method not founded on the work of Lilhanand and Tseng (1988). We employ a signal-processing approach that models a general nonlinear transformation of one signal into another. This makes use of state-of-the-art nonlinear optimization and wavelet decomposition. We seek an algorithm that is stable (i.e., it always converges to some useful solution) and robust (i.e., it is insensitive to target response spectral shape or seed record magnitude). We seek a modified signal with an excellent match to the target spectrum, while maintaining the original seed signal’s nonstationary statistics, such as general envelope, time location of large pulses, and variation of frequency content with time.

Theory

The goal here is to develop a process that transforms a known signal $x(t)$ into a similar but spectrum-matching one, $y(t)$. An additional aim is that the transformed signal $y(t)$ should maintain qualitatively similar nonstationary characteristics, such as general envelope, time location of large pulses, and variation of frequency content with time. Thus, we seek a signal $y(t)$ that looks, for all intents and purposes, like a real earthquake. A list of symbols used in this paper is presented in Table 1.

Application of Volterra Series

The Volterra series originates as a generalization of the Taylor series expansion of a function. However, it can also be considered as a generalization of the convolution integral, a transformation of an input function in time by some system. A reliable method of selecting an optimal Volterra model is not generally available. Therefore, any problem, whether
Obtaining Spectrum Matching Time Series Using RVSA
1665

linear or nonlinear, is unique and particular to its circumstances. So the choice of the Volterra model is, in a sense, arbitrary. In this paper, the wavelet transform is employed in order to estimate the Volterra kernels of the series expansion (of \( x(t) \)) that maps \( x(t) \) into an infinite set of \( y(t) \). Given that \( y(t) \) is some Volterra series expansion of \( x(t) \), to some degree it inherits the shape and form of \( x(t) \). Consequently, from this infinite set of \( y(t) \), we seek a particular member that has a response spectrum that matches some target response spectrum.

Consider some nonlinear process that transforms a frequency domain signal \( X(\omega) \) into \( Y(\omega) \). The general form of this process is given by a Volterra series expansion (Schetzen, 1980; Boyd et al., 1984; Chanerley et al., 2007; Fa-Long, 2011):

\[
Y(\omega) = \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int H_n(\omega_1, \cdots, \omega_n) X(\omega - \omega_1) \cdots X(\omega - \omega_n) d\omega_1 \cdots d\omega_n,
\]

in which \( H_n \) is known as the \( n \)th Volterra kernel. This can be inverse Fourier transformed into the time domain using

\[
y(t) = \sum_{n=1}^{\infty} \frac{1}{n!} h_n(t)x(t)^n,
\]

in which \( y(t) \) and \( x(t) \) are the inverse Fourier transforms of signals \( Y(\omega) \) and \( X(\omega) \). The time-domain version of the \( n \)th Volterra kernel \( H_n \) is \( h_n \). This is similar to a Taylor series, but the coefficients are themselves time-varying functions.

Let original signal \( x(t) \) be re-expressed as the linear combination of narrowband frequency components \( \phi_i(t) \),

\[
x(t) \approx \sum_{i=1}^{m} a_i \phi_i(t),
\]

in which \( a_i \) are amplitude coefficients. This decomposition could be a Fourier series, but in this paper we shall employ a wavelet decomposition using the stationary wavelet transform (SWT; Chanerley and Alexander, 2002, 2007, 2010; Berrill et al., 2011; Chanerley et al., 2013). In expression (3), the total number of terms is \( m \), that is \( m \) = 1 wavelet detail levels plus 1 wavelet approximation level. Hence, technically, the number of levels of wavelet decomposition should be considered \( m-1 \) rather than \( m \).

The more commonly used transform is the discrete wavelet transform (DWT), which operates using digital low-pass and high-pass filters and then decimates the data, but decimation leads to shift variance and aliasing. In effect, the DWT are octave digital filters. However, the filter banks employed in the DWT will experience aliasing and phase and amplitude distortion. The filters are usually overlapping, therefore aliasing errors will inevitably occur. It can be shown that the reconstructed signal will be aliased as a consequence of decimation and imaging. Moreover, any processing is performed only on the data available after downsampling, therefore missing data samples are unprocessed. Furthermore, during reconstruction the interpolation will proceed on the available processed samples resulting in some distortion.

Therefore, we turn to the SWT, also known as the translation invariant transform, which operates in a different manner to that of the standard DWT. The SWT is shift invariant and does not alias. In this case, the number of data samples is kept the same because downsampling is avoided at each level of decomposition. Instead, the filter impulse response is interpolated to match the filter bandwidths to that of the subbands. The interpolation is performed by inserting zeros between the filter coefficients; this method is referred to as the "a trous" algorithm ("with holes" algorithm). During reconstruction, it is not now necessary to upsample the data such as when applying the DWT. However, it is necessary to apply synthesizing filters for the DWT, which are flipped versions of the analysis filters at each decomposition level.

In order to maintain a linear phase characteristic, the biorthogonal wavelet family is recommended and the higher-order bior3.9 wavelet (MATLAB wavelet toolbox, 2010; see Data and Resources) was used to ensure the appropriate filter characteristics for wavelet approximation and detail estimation.

Substituting equation (3) into equation (2), we obtain

\[
y(t) = \sum_{i=1}^{m} [h_1(t) a_i] \phi_i(t) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} [h_2(t) a_i a_j] \phi_i(t) \phi_j(t) + \cdots
\]

or, in the frequency domain,

\[
Y(\omega) = \sum_{i=1}^{m} \int [H_1(\omega) a_i] \Phi_i(\omega - \omega_1) d\omega_1 + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \int [H_2(\omega_1, \omega_2) a_i a_j] \Phi_i(\omega - \omega_1) \Phi_j(\omega - \omega_2) d\omega_1 d\omega_2 + \cdots.
\]

If we assume \( \phi_i(t) \) is a narrowband frequency component and that Volterra kernels are almost constant across this narrowband function, then the following approximation is reasonable:

\[
H_1 a_i = \beta_i, \quad H_2 a_i a_j = \beta_{ij}, \quad H_3 a_i a_j a_k = \beta_{ijk}, \ldots,
\]

in which coefficients \( \beta_i, \beta_{ij}, \beta_{ijk}, \ldots \) are all frequency/time-invariant weights. In this case, these coefficients can be considered to represent a discrete approximation of the continuous Volterra kernels. Hence, we re-express the general Volterra series, equation (4), as the following linear combination of \( \phi_i(t) \) and all positive multinomial combinations of \( \phi_i(t) \):

\[
y(t) = \beta^T \psi(t),
\]

in which vectors \( \beta \) and \( \psi \) are given as
in which subscripts \( i, j, k \), etc. are members of the set of integer from 1 to \( m \). This Volterra series expansion (e.g., \( y(t) \) in equation 7) is expressed simply as linear combinations of functions within \( \psi \). It is worth considering the implication of expressions (7) and (8). In the case, where a wavelet vector basis equation (3) is employed, it is clear that each \( \phi_i \) shares a similar nonstationary envelope of its progenitor signal \( x(t) \). Hence all multinomial combinations (i.e., \( \phi_1 \phi_k, \phi_1 \phi_j \phi_k \), and so on) also share this similar nonstationary envelope. Figure 1 graphically displays an example of the Volterra basis \( \psi(t) \) for the case of one wavelet level of decomposition. We use only one wavelet level here for pictorial clarity. For all other results in the paper, the maximum number of wavelet levels is adopted, which is typically in the range of 8–12 levels (for recorded accelerograms). The exact number is dependent on the mother wavelet and seed record length (see the \texttt{wmaxlev} function in MATLAB; see Data and Resources). In this figure the detail (higher-frequency subband) is \( \phi_1(t) \), and the approximation (the lower-frequency subband) is \( \phi_2(t) \). These functions, when multiplied by weights \( \beta \) and summed, form the first Volterra kernel. All weighted and summed quadratic and cubic combinations of these functions form the second and third Volterra kernels.

Thus, reweighting (i.e., changing) the \( \beta \) coefficients in equation (7) allows for a modification of frequency and phase content of the signal while keeping the general form of the nonstationary nature of the signal \( x(t) \). That is to say \( x(t) \) and \( y(t) \) can have qualitatively similar nonstationary features (e.g., very similar envelopes and pulses).

The general form of the modified time series \( y(t) \) (in equation 7) is simply a weighted sum of wavelet levels \( \phi_1(t) \) of \( x(t) \) (i.e., the first Volterra kernel), plus the weighted sum of quadratic combinations of wavelet levels \( \phi_1(t) \phi_2(t) \) (the second Volterra kernel), plus the weighted sum of cubic combinations of wavelet levels \( \phi_1(t) \phi_2(t) \phi_k(t) \) (the third Volterra kernel), etc. By adjusting the weights, for example, \( \beta_1, \beta_2, \) and \( \beta_{jk} \), we obtain an infinite set of time series that are related to the original time series \( x(t) \) by some nonlinear transformation. Thus, we are searching this infinite set for a particular member of \( y(t) \) that has the property that its response spectrum approximately matches some target response spectrum.

The size \( q \) of the Volterra weight vector \( \beta \) and basis vector \( \psi \) is given by the summed number of terms in a set of multinomial expansions, which can be shown as

\[
q = 1 + \frac{(p + m)!}{m! (m - 1)!} - 1,
\]

in which \( p \) is the highest order of Volterra kernels employed and \( m - 1 \) is the number wavelet decomposition levels of the original accelerogram.

Nonlinear Least Squares/Norm Problem

Now consider a response spectrum \( s(y(t), f) \) produced from a signal \( y(t) \) at a discrete set of \( r \) frequencies in vector \( f \). Here, the original signal \( x(t) \) and morphed signal \( y(t) \) are accelerograms. Evaluating the function \( s(y(t), f) \) involves the numerical solution of a second-order differential system for a single-degree-of-freedom system. Many integration schemes exist in the research literature. For example, Newmark’s method (see Clough and Penzian, 1993), the fast Fourier transform, and the Smallwood (1981) filter (that is based on the Laplace transform) can all be employed. However, Smallwood’s time-domain convolution filter is by far the quickest algorithm and is reasonably accurate.
Thus, the problem of morphing a real accelerogram $x(t)$ into a spectrum-compatible one $y(t)$ can now be expressed as a nonlinear least squares/norm problem. We seek to minimize the difference between the spectrum $s(y(t), f)$ and some target spectrum $s_T(f)$ by selecting the optimal and re-weighted coefficients $\beta$.

An objective function $v(\beta)$ for this nonlinear least squares/norm problem can be stated as

$$\min_{\beta} \|v(\beta)\|, \quad \text{in which } v(\beta) = \frac{s(y(t), f) - s_T(f)}{s_T(f)}. \quad (10)$$

Equation (10) represents $r$ nonlinear algebraic equations with $q$ unknown coefficients $\beta$. We seek to minimize the Euclidian norm of $v(\beta)$ in order to determine the optimal spectrum-matching morphed $y(t)$. So the morphed time series is based on optimal and reweighted coefficients $\beta$. This problem can be solved using the Levenberg–Marquardt algorithm (Moré, 1977; Shaterenikht and Alexander, 2012). This algorithm is preferable to the trust-region-reflective approach (Coleman and Li, 1996) as it is able to deal with both the overdetermined (least square) problem (i.e., $q < r$) and underdetermined ones (least norm; i.e., $q > r$).

The solutions obtained from the Levenberg–Marquardt algorithm to this nonlinear optimization problem are only local optima and not a guaranteed global optimum. The searching algorithm that would be required for a global optimum is extremely costly in terms of computational resources (i.e., time and memory), thus it has not been explored in this paper. We present only the first local optimum obtained from the given initial start and show that it is sufficiently good for our purposes. The initial start we have employed is the original record $x(t)$, which is defined by the following coefficients $\beta_i = 1$, $i \in [1, m]$, and all other coefficients $\beta_i^j = 0$. That is to say, initially, all coefficients of first-order Volterra kernel are one, and all coefficients for higher-order Volterra kernels are zero. Thus, at the start of the optimization process, $y(t)$ equals $x(t)$.

In summary, we define a nonlinear process that transforms a time series $x(t)$ into $y(t)$. This general process is defined in terms of a Volterra series expansion of $x(t)$. From this general expansion, we search for a particular case in which $y(t)$ has a response spectrum that approximately matches some target response spectrum. This particular $y(t)$ is determined by adjusting the weights (coefficients) of the Volterra series using nonlinear optimization.

### Exploring the Performance of the Reweighted Volterra Series Algorithm (RVSA)

#### Selection of Seed Records

Bray and Travasarou (2007), Athanasopoulos (2008), and Rathje et al. (2010) show that the selection process of ground motion can have a large effect on the goodness of fit of a spectral matching process. With this in mind, we select records from the PEER-NGA ground-motion database (PEER, 2010). The records are selected so that their geometric mean is a spectral match to the target spectrum. The additional stratified selection criteria are that the records should be from events of magnitude 6–7.5 and should be from free field sites of shear-wave velocity $V_s \geq 800$ m/s. The scaling of these records is performed by PEER-NGA such that the geometric mean of this record set is a reasonable match to the target spectrum. The dataset, target spectrum, and geometric mean of records are displayed in Figure 2. The philosophy

---

**Table 1**

List of Commonly Used Mathematical Symbols in this Paper and their Definitions

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_i$</td>
<td>Amplitude coefficients of original signal $x(t)$</td>
</tr>
<tr>
<td>$f$</td>
<td>Vector of frequencies (size $r$)</td>
</tr>
<tr>
<td>$h_n$</td>
<td>The time-domain versions of $n$th Volterra kernel</td>
</tr>
<tr>
<td>$H_n$</td>
<td>The frequency-domain versions of $n$th Volterra kernel</td>
</tr>
<tr>
<td>$m-1$</td>
<td>Number of wavelet decomposition levels of original accelerogram</td>
</tr>
<tr>
<td>$p$</td>
<td>Highest Volterra kernel employed</td>
</tr>
<tr>
<td>$q$</td>
<td>Total number of terms in Volterra series expansion (number of unknowns)</td>
</tr>
<tr>
<td>$r$</td>
<td>Total number of points that are spectrally matched (number of equations)</td>
</tr>
<tr>
<td>$s(y, f)$</td>
<td>Response spectrum of signal $y(t)$ at frequency points $f$</td>
</tr>
<tr>
<td>$s_T(f)$</td>
<td>Target response spectrum at frequency points $f$</td>
</tr>
<tr>
<td>$v(\beta)$</td>
<td>Objective vector function (size $r$) for minimization</td>
</tr>
<tr>
<td>$x(t)$</td>
<td>Time-domain version of original signal</td>
</tr>
<tr>
<td>$X(\omega)$</td>
<td>Frequency-domain version of original signal</td>
</tr>
<tr>
<td>$y(t)$</td>
<td>Time-domain version of morphed, spectrum matching, signal</td>
</tr>
<tr>
<td>$Y(\omega)$</td>
<td>Frequency-domain version of morphed, spectrum matching, signal</td>
</tr>
<tr>
<td>$\beta_i$</td>
<td>Weights/coefficients of Volterra series for morphed signal $y(t)$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Vector of weights/coefficients (size $q$) of Volterra series for morphed signal $y(t)$</td>
</tr>
<tr>
<td>$\phi_i(t)$</td>
<td>$i$th level wavelet decomposition (in time-domain) of signal $x(t)$</td>
</tr>
<tr>
<td>$\Phi_i(\omega)$</td>
<td>$i$th level wavelet decomposition (in frequency-domain) of signal $x(t)$</td>
</tr>
<tr>
<td>$\psi(t)$</td>
<td>Vector of functions (basis) that are used to define the Volterra kernels (size $q$).</td>
</tr>
</tbody>
</table>

---

**Figure 2.** PEER NGA data set used in this paper.
here is that the seed records are close in magnitude to the target spectrum. A list of records selected is given in Table 2.

Influence of Number of Volterra Kernels, \( p \)

As a heuristic case, a Eurocode 8 (2004) type I horizontal elastic spectrum of \( 0.35 g \), on soil class A, with 5\% of critical damping, is employed in this paper. This total acceleration response spectrum was defined from 0.02 to 4 s using 200 divisions on a logarithmic scale (i.e., \( r = 200 \)).

Table 3 displays a summary comparison of the performance to the method proposed in this paper for the records in Table 2. As the number of Volterra kernels is increased, the match to the spectrum improves. When three kernels are included, the mean error (of 20 records, for 3 Volterra kernels) is 0.04\%, with a mean maximal error of 7.81\%. These values are clearly very small and good.

Results show that it is of marginal value to employ more than three Volterra kernels. This is because (1) the problem is most likely underdetermined and (2) the total number of Volterra terms \( q \) grows exponentially, and this growth in the problem size gives rise to problems of computational efficiency and computability.

Comparison with SeismoMatch (see Table 3)

Table 3 describes the performance of SeismoMatch and RVSA. The RVSA is found to be superior in terms of the quality of the match and robustness of the matching process. SeismoMatch converged well for 6 of the 20 records, the “f” flag in the table signifying the nonconvergence of the algorithm.

Table 2

<table>
<thead>
<tr>
<th>NGA Record Number</th>
<th>Scale Factor</th>
<th>Event</th>
<th>Year</th>
<th>Station</th>
<th>Moment Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>284</td>
<td>5.8166</td>
<td>Irpinia-Italy-01</td>
<td>1980</td>
<td>Auletta</td>
<td>6.9</td>
</tr>
<tr>
<td>285</td>
<td>1.765</td>
<td>Irpinia-Italy-01</td>
<td>1980</td>
<td>Bagnolo Irpinio</td>
<td>6.9</td>
</tr>
<tr>
<td>292</td>
<td>0.9085</td>
<td>Irpinia-Italy-01</td>
<td>1980</td>
<td>Sturno</td>
<td>6.9</td>
</tr>
<tr>
<td>296</td>
<td>8.5355</td>
<td>Irpinia-Italy-02</td>
<td>1980</td>
<td>Bagnolo Irpinio</td>
<td>6.2</td>
</tr>
<tr>
<td>297</td>
<td>3.8295</td>
<td>Irpinia-Italy-02</td>
<td>1980</td>
<td>Bisaccia</td>
<td>6.2</td>
</tr>
<tr>
<td>303</td>
<td>5.8806</td>
<td>Irpinia-Italy-02</td>
<td>1980</td>
<td>Sturno</td>
<td>6.2</td>
</tr>
<tr>
<td>455</td>
<td>9.3573</td>
<td>Morgan Hill</td>
<td>1984</td>
<td>Gilroy Array Number 1</td>
<td>6.19</td>
</tr>
<tr>
<td>765</td>
<td>1.1548</td>
<td>Loma Prieta</td>
<td>1989</td>
<td>Gilroy Array Number 1</td>
<td>6.93</td>
</tr>
<tr>
<td>957</td>
<td>3.3283</td>
<td>Northridge-01</td>
<td>1994</td>
<td>Burbank–Howard Road</td>
<td>6.69</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>NGA Record Number</th>
<th>SeismoMatch v.2 Mean Error (%)</th>
<th>SeismoMatch v.2 Max Error (%)</th>
<th>RVSA 1 Kernel Mean Error (%)</th>
<th>RVSA 1 Kernel Max Error (%)</th>
<th>RVSA 2 Kernels Mean Error (%)</th>
<th>RVSA 2 Kernels Max Error (%)</th>
<th>RVSA 3 Kernels Mean Error (%)</th>
<th>RVSA 3 Kernels Max Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>284FN</td>
<td>13.30</td>
<td>58.40</td>
<td>−2.97</td>
<td>48.01</td>
<td>−0.18</td>
<td>17.93</td>
<td>−0.05</td>
<td>7.33</td>
</tr>
<tr>
<td>284FP</td>
<td>5.50</td>
<td>28.70</td>
<td>−1.92</td>
<td>37.92</td>
<td>−0.37</td>
<td>17.47</td>
<td>−0.10</td>
<td>14.53</td>
</tr>
<tr>
<td>285FN</td>
<td>5.90</td>
<td>29.90</td>
<td>−1.20</td>
<td>35.03</td>
<td>−0.22</td>
<td>17.75</td>
<td>−0.01</td>
<td>3.80</td>
</tr>
<tr>
<td>285FP</td>
<td>6.90</td>
<td>29.30</td>
<td>−1.20</td>
<td>35.03</td>
<td>−0.22</td>
<td>17.75</td>
<td>−0.01</td>
<td>3.80</td>
</tr>
<tr>
<td>292FN</td>
<td>10.50</td>
<td>46.50</td>
<td>−1.64</td>
<td>41.99</td>
<td>−0.22</td>
<td>17.75</td>
<td>−0.01</td>
<td>3.80</td>
</tr>
<tr>
<td>292FP</td>
<td>7.20</td>
<td>39.80</td>
<td>−1.30</td>
<td>26.74</td>
<td>−0.18</td>
<td>12.54</td>
<td>−0.01</td>
<td>3.49</td>
</tr>
<tr>
<td>296FN</td>
<td>7.70</td>
<td>29.70</td>
<td>−1.48</td>
<td>36.99</td>
<td>−0.23</td>
<td>15.96</td>
<td>−0.11</td>
<td>19.11</td>
</tr>
<tr>
<td>296FP</td>
<td>7.50</td>
<td>32.50</td>
<td>−2.15</td>
<td>45.35</td>
<td>−0.12</td>
<td>10.70</td>
<td>−0.01</td>
<td>3.27</td>
</tr>
<tr>
<td>297FN</td>
<td>12.10</td>
<td>53.50</td>
<td>−2.98</td>
<td>53.24</td>
<td>−0.24</td>
<td>21.61</td>
<td>−0.01</td>
<td>3.46</td>
</tr>
<tr>
<td>297FP</td>
<td>10.80</td>
<td>46.20</td>
<td>−2.72</td>
<td>41.50</td>
<td>−0.17</td>
<td>14.39</td>
<td>−0.07</td>
<td>10.59</td>
</tr>
<tr>
<td>303FN</td>
<td>10.90</td>
<td>62.00</td>
<td>−2.23</td>
<td>46.94</td>
<td>−0.38</td>
<td>18.79</td>
<td>−0.05</td>
<td>10.26</td>
</tr>
<tr>
<td>303FP</td>
<td>9.10</td>
<td>51.40</td>
<td>−3.03</td>
<td>55.99</td>
<td>−0.18</td>
<td>19.10</td>
<td>−0.03</td>
<td>9.40</td>
</tr>
<tr>
<td>455FN</td>
<td>18.80</td>
<td>356.80</td>
<td>−3.07</td>
<td>41.74</td>
<td>−0.19</td>
<td>14.82</td>
<td>−0.07</td>
<td>11.61</td>
</tr>
<tr>
<td>455FP</td>
<td>26.80</td>
<td>222.00</td>
<td>−1.87</td>
<td>33.93</td>
<td>−0.17</td>
<td>16.72</td>
<td>−0.08</td>
<td>13.83</td>
</tr>
<tr>
<td>765FN</td>
<td>10.70</td>
<td>34.90</td>
<td>−2.06</td>
<td>42.94</td>
<td>−0.10</td>
<td>17.14</td>
<td>−0.01</td>
<td>3.91</td>
</tr>
<tr>
<td>765FP</td>
<td>11.80</td>
<td>29.90</td>
<td>−2.40</td>
<td>38.94</td>
<td>−0.06</td>
<td>10.54</td>
<td>−0.01</td>
<td>3.05</td>
</tr>
<tr>
<td>957FN</td>
<td>13.30</td>
<td>58.00</td>
<td>−1.66</td>
<td>39.43</td>
<td>−0.10</td>
<td>12.45</td>
<td>−0.04</td>
<td>8.60</td>
</tr>
<tr>
<td>957FP</td>
<td>14.00</td>
<td>132.60</td>
<td>−1.34</td>
<td>36.00</td>
<td>−0.24</td>
<td>19.46</td>
<td>−0.04</td>
<td>7.89</td>
</tr>
<tr>
<td>1011FN</td>
<td>5.70</td>
<td>27.80</td>
<td>−1.03</td>
<td>29.25</td>
<td>−0.05</td>
<td>7.87</td>
<td>−0.02</td>
<td>6.40</td>
</tr>
<tr>
<td>1011FP</td>
<td>10.00</td>
<td>70.30</td>
<td>−1.90</td>
<td>42.60</td>
<td>−0.27</td>
<td>15.54</td>
<td>−0.02</td>
<td>6.81</td>
</tr>
<tr>
<td>Mean</td>
<td>7.25</td>
<td>29.22</td>
<td>−2.01</td>
<td>40.48</td>
<td>−0.19</td>
<td>15.60</td>
<td>−0.04</td>
<td>7.81</td>
</tr>
</tbody>
</table>
The mean error of the 6 that converged was 7.25%, with a mean maximal error of 29.22%. In addition, SeismoMatch is quite sensitive to the amplitude of the seed accelerograms. Scaling a convergent record by a different amount can result in nonconvergence. This requires careful selection of records and, to some degree, a trial-and-error approach. In contrast, RVSA is insensitive to input acceleration amplitudes. It can produce a match even if the seed record is not scaled so as to be an approximate match to the target spectrum.

Consider Figure 3, which shows the match for record NGA 285 (fault normal [FN]). This seed accelerogram was selected as it is one of the best performing SeismoMatch records observed. Figure 3a displays how the goodness of the match improves with the number of Volterra kernels and indicates the match with RVSA (when three Volterra kernels are used) is very good.

Yet it is worth considering the accelerogram time histories in Figure 4. Figure 4a depicts the performance of RVSA, and Figure 4b shows the performance of SeismoMatch. Both SeismoMatch and RVSA conserve much of the original seed record, maintaining similar nonstationary statistics.

Figure 5 displays the response spectra and time histories for the best- and worst-performing cases displayed of RVSA in Table 3. This figure indicates the spectral match (using three Volterra kernels) is generally very good.

Reviewing Energy Content of Morphed and Original Records

Figure 6 shows an example of the normalized cumulative energy for both the original and spectrum-matched time series (for seed record NGA 296 FN). The temporal distribution of energy is very similar in both cases.

Time–frequency analysis enables a more forensic comparison of original and morphed records. In this paper, we use the time–frequency toolbox (Auger et al., 1996), which uses a smoothed pseudo-Wigner–Ville approach. Figure 7 displays the time–frequency plot of the original signal NGA 296 (FN). This plot indicates the key power occurs from 6 to 8 s and at around 2 Hz. Contrast this with Figure 8, which displays the spectrum-matching NGA 296 (FN) record. More power is added between 2 and 5 Hz. Thus the matched record has a slightly broader range of high-amplitude frequency components, as is to be expected. For both the original NGA 296 and the spectrum-matched NGA 296, there is no significant power above 9 Hz.

The added power \( \frac{y(t)}{x(t)} \) is at time locations where some moderately significant power is present in the original signal \( x(t) \). In this case, the additional power is added in the 6–8 s time range so as to not disrupt the overall envelope or location of large pulses, etc. Thus, we can see that the nonstationary characteristics of original \( x(t) \) and spectrum-matched \( y(t) \) records are qualitative comparable.

Robustness to Noise in Seed Signals

One could ask the question: how sensitive is the algorithm to noisy input signals \( x(t) \)? This is because many corrected records have been filtered. Frequency components of the time series in the stop bands represent very low-power noisy signals. RVSA, in its nonlinear optimization, may attempt to recover the noisy filtered-out terms rather like
Chanerley and Alexander (2007, 2008). Thus, the question arises as to how sensitive this processes is to signal noise. As a worst-case scenario, we consider a completely artificial Gaussian signal that is enveloped in time (i.e., the whole signal is noise). What is the performance of the proposed algorithm with a signal like this? Figure 9 displays the results of using a noisy input signal. The Volterra series method still manages to produce a spectrum-compatible time series that keeps its general envelope. The mean error is 0.08%, with a maximal error of 12.4%. This error is of the same order as that obtained with real accelerograms as seeds. Thus, we conclude that the proposed algorithm is fairly robust to noise in the input signal. However, it will use whatever information is present in the seed time series, whether it is signal or noise.

Stability of Integrated Acceleration Time Series

The previous analyses have demonstrated the efficacy of using RVSA for modifying a recorded accelerogram into a spectrum-matching one. However, it is important to review the stability and plausibility of the first and second integral of the acceleration, that is the velocity and displacement time series. The question of numerically integrating accelerograms is a problematic one (Graizer, 2010). Integration is fundamentally a low-pass filter that attenuates the high-frequency content and amplifies the very low-frequency content of a signal (Chanerley and Alexander, 2010). The very low-frequency content of an accelerogram is frequently corrupted by noise and ground pitching and rolling degrees of freedom (often described as tilts) (Graizer, 2005). Regularly, as in the PEER-NGA database, this low-frequency content is removed by filtering before and possibly after integration (Converse, 1992) to remove these troublesome artifacts. However, Chanerley and Alexander (2010) and Chanerley et al. (2013) pointed out that this can remove very-low-frequency fling components of the ground motion. Hence, the resultant

Figure 5. Response spectra and time histories of best- and worst-performing cases.

Figure 6. Cumulative energy versus time for original and spectrum-matched record, NGA 296 fault normal (FN).
ground-motion displacement time series is stable but erroneous in its low frequencies.

Hence, with spectrum-matched records produced by RVSA, we could (1) adopt the algorithm in Chanerley and Alexander (2010) and Chanerley et al. (2013) or (2) integrate and filter out the low-frequency components. In this paper, all records have been processed and corrected by PEER and have already been low-cut filtered, thus, in this case, adopting option (1) is not consistent. Instead, we filter the matched accelerogram below 0.2 Hz using a fourth-order zero-phase Butterworth filter before and after the integration to obtain stable velocity and displacement time series. An example of these is shown in Figure 10.

Conclusions

In this paper, we present a novel algorithm (based on state-of-the-art signal processing and optimization techniques) for modifying a given ground acceleration time series such that its response spectrum matches a given target response spectrum.

The RVSA demonstrates stability and robustness. Unlike some approaches, it generally appears to converge to some useful record that meets the objectives of the spectral matching process. It can be used regardless of which code-based spectral shape is chosen as a target spectrum. In addition, it can be classed as insensitive to the seed record selected (i.e., its magnitude or noise level). That is to say, it converges to any given and plausible response spectrum from any given and plausible seed time series. This includes enveloped Gaussian noise as a given seed time series.

The spectrum-matching time series maintains a very similar nonstationary characteristic (e.g., its general envelope, time location of large pulses, and variation of frequency content with time) to the seed time series. It is worth stating that the matched time series has the same total length as the seed record. RVSA does not affect this length in the matching...

Figure 7. Time–frequency plot of original NGA 296 FN using the pseudo-Wigner–Ville method. The color version of this figure is available only in the electronic edition.

Figure 8. Time–frequency plot of spectrum matching NGA 296 FN (using RVSA) using the pseudo-Wigner–Ville method. The color version of this figure is available only in the electronic edition.

Figure 9. Example of algorithm with time-enveloped Gaussian noise N(0,1) as seed.
processes. Thus, if duration matching is an additional requirement, then it is sensible to select an appropriately long seed record. For the case of real accelerogram seeds, the resultant spectrum-matched record visually appears to be like an actual recorded ground-motion time series.

As the number of Volterra kernels employed increases up to three, we observe an increase in the goodness of fit to the spectrum at the expense of computational time. Employing more than three kernels was found to be ineffective, as the problem was more likely to become underdetermined. In this study, we show that, at three kernels, the mean misfit error (of the spectral match) was 0.04%, with a mean maximal misfit error of 7.8% (at any frequency point of that spectrum). Thus, spectrum-matched time series show an excellent fit to the target spectrum over the entire structural frequency range (and much better than was previously available), while maintaining a record that keeps a qualitatively similar appearance to the seed record.

**Data and Resources**

All strong-motion data for seed records were obtained from Pacific Earthquake Engineering Research Center (PEER) (2010; data last accessed 14 January 2013). The reweighted volterra series algorithm (RVSA) code (MATLAB) is available by e-mail from the author, N.A.A.

**Acknowledgments**

The authors wish to thank the Institute of Advanced Studies, at the University of Bristol, and the universities of East London and Surrey for every support given.

**References**


Japanese Society of Civil Engineers (1997). Dynamic analysis and earthquake resistant design, in Strong Motion and Dynamic Properties, A.A. Balkema, Rotterdam, Netherlands.


Teldorina Technologies (2002). Network Equipment-Building System (NEBS) requirements: Physical protection, Teldorina Technologies GR-63-CORE.

Department of Civil Engineering
Room 2.33
Queen’s Building
Bristol BS8 1TR, United Kingdom
nick.alexander@bris.ac.uk
(N.A.A., A.J.C.)

University of East London
School of Architecture, Computing & Engineering
Dockland Campus
London E16 2RD United Kingdom
(A.A.C.)

Department of Civil and Environmental Engineering
Faculty of Engineering and Physical Sciences
University of Surrey
Thomas Telford Building
Guildford, Surrey GU2 7XH
(S.B.)

Manuscript received 26 July 2013;
Published Online 15 July 2014