Originally published as:


DOI: http://doi.org/10.1785/0120150281
Bulletin of the Seismological Society of America

This copy is for distribution only by the authors of the article and their institutions in accordance with the Open Access Policy of the Seismological Society of America.

For more information see the publications section of the SSA website at www.seismosoc.org
Ground-Motion Forecasting Using a Reference Station and Complex Site-Response Functions Accounting for the Shallow Geology

by Marco Pilz* and Stefano Parolai

Abstract The distribution of damage due to recent earthquakes has shown that the effects of shallow geological structures on the level of ground shaking represent an important factor in engineering seismology. Whereas many previous studies have estimated site amplification factors in the frequency domain, their application to the real-time modeling of ground motion is not yet fully established. In this article, a method for the real-time correction of frequency-dependent site-response factors is proposed, which accounts not only for the modulus, but also for the changes in the signal phase related to local site conditions. The transformation of the complex standard spectral ratios to a causal recursive filter in the time domain allows for the forecasting of the waveforms for soft-soil sites almost in real time when the signal is recorded earlier at a reference site. When considering travel-time differences of the various seismic phases between the hypocenter and the studied sites, the level of ground motion at soft-soil sites with respect to arrival time, energy, duration, and frequency content can be well constrained, even in cases of a high spatial variability of the amplification patterns.

Introduction

The development of mathematical models for the simulation of ground-motion time series for arbitrary sites, starting from the recordings at a reference site, has always been an important issue for any kind of seismic hazard and risk assessment. In particular, this task might become an important issue for regional and onsite earthquake early warning and rapid response systems. When earthquake ground motion is recorded at one specific site, it is influenced by many factors, such as source mechanism, propagation path of the seismic waves, as well as local site conditions that may involve complex surface geology and irregular topography. The amplitudes of seismic waves and the duration of shaking are significantly modified as the seismic waves propagate through the soft-soil layers near the surface. This phenomenon is known as site amplification, and it significantly influences the damage to structures. In addition to the level of ground motion, experience from a number of earthquakes has shown that little damage to structures occurs in cases where the ground shaking has a short duration, even though the accelerations and spectral amplitudes may be large (e.g., Housner and Trifunac, 1967; Dobry et al., 1978; Bommer and Abrahamson, 2006). This means that seismic waveforms recorded at two adjacent stations, even if they are very far away from the hypocenter, may be very different in amplitude and duration if their site amplification factors are very different.

Starting already more than half a century ago, a large number of theoretical and experimental investigations, as well as analytical and empirical models of site amplification, have been published (e.g., Kanai, 1952; Gutenberg, 1957; Idriss and Seed, 1968; Borcherdt, 1970, 1994; Roesset, 1970; Mohraz, 1976; Phillips and Aki, 1986; Seed et al., 1988; Şafak, 1989; Boore and Joyner, 1991). Concurrent with the accurate simulation of seismic site response, which started in the 1970s, including programs for the propagation of seismic waves through a 1D soil column such as equivalent linear seismic response analysis of horizontally layered soil deposits (SHAKE; Schnabel et al., 1972), dynamic effective stress response analysis of soil deposits (DESRA-2; Lee and Finn, 1975), and characteristics method applied to soils (CHARSOIL; Streeter et al., 1974), relevant aspects of the 1D (and 2D) wave propagation problems have been extensively examined by seismologists.

A popular approach for studying the dynamic site response of a 1D soil column is the approximation through a single-degree-of-freedom (SDOF) system (Roesset et al., 1994). In general, methods for analyzing the dynamic response of SDOF systems can be based in both the time and frequency domains. For the latter case, the seismic response can be gained by the Fourier integral transform and inverse transform techniques, provided that the accurate site-response function is known. On the other hand, time-domain methods aim at calculating a causal digital filter so as to make the site-response function close to the theoretical transfer function.

*Now at Swiss Seismological Service, ETH Zürich, Sonneggstr. 5, NO H 60, 8092 Zürich, Switzerland.
in the frequency range that engineers are interested in. The
time-domain solution has three attractive attributes: (a) comput-
tional problems such as frequency-domain aliasing can be
avoided, (b) time-domain methods are capable of directly
incorporating nonlinear and time-varying behaviors, and (c) the
time-domain approach preserves causality and main-
tains the arrow of time.

For problems in seismology, several ways have been pro-
posed for the optimal calculation of site-response functions
(an overview is given by Ólafsson and Sighjörnsdóttir, 2011).
For example, auto-regressive (AR) algorithms for wave motion
were developed by Jurkevics and Ulrych (1978, 1979) with
follow-up works on auto-regressive moving-average (ARMA)
algorithms by Polhemus and Cakmak (1981), Chan and Parks
(1982), Nau et al. (1982), and Spanos (1983). Recursive sim-
ulations of multivariate random time series were presented by
records was pioneered by Kozin (1988). Conte et al. (1992)
applied an ARMA model to simulate strong ground motion
using the Kanai–Tajimi spectral representation for ground
acceleration (Kanai, 1957; Tajimi, 1960). Mignolet and
Spanos (1991) used AR, MA (moving average), and ARMA
models to simulate 2D random fields. Kanamori et al. (1999),
using the forward difference method, derived the recursive
digital filter for the continuous processing of various ground-
motion parameters. Lee (1984), using the Z-transform tech-
nique, studied the same oscillator and presented recursive
formulas of the relative displacement, relative velocity, and
relative acceleration. Beck and Dowling (1988), using Duha-
mel’s step-integral method and assuming successive linear
segments of input acceleration, concluded that the Lee (1984)
method does not give an exact solution for arbitrary forcing
functions as in the case of earthquake records. Remarkably,
Ólafsson et al. (2001) noted that ARMA models can capture
the most essential features of earthquake ground motion.
However, they also concluded that such models might lack
consideration of the frequency content of ground motion which
might be particularly important for systems with degrading properties. Kanamori and Rivera (2008)
used a time-domain filter similar to Zhu (2003) for near-real-
time moment tensor inversion using the W-phase for rapid
 tsunami warnings. Combining the relative amplification with
 propagation factor, Nagashima et al. (2008) presented a
method for the prediction of low frequency by applying a filter
to the waveforms observed at front site. Kuyuk and Motosaka
(2008, 2009) proposed a new method to predict waveforms at
a target site in real time using the waveforms at front site based
on the artificial-neural-network methodology. Maeda et al.
(2011), using recordings of the dense Hi-net seismographic
network in Japan, designed a recursive time-domain filter that
recovers interference effects of long-period seismic wavefields
with large, distant earthquakes. Recently, Hoshiba (2013)
presented a method for the correction of site amplification factors
by applying an infinite impulse response filter to correct the
site factor in real time.

However, so far the focus has been on modeling the am-
plitude of the site-response function ignoring phase charac-
teristics, whereas for site-specific modeling, the increase of
amplification related to peculiar 2D (valley) or 3D (basin)
geometries is still only very poorly quantified. After the long
(and not yet resolved) debate about the relative importance of
such 3D effects in Mexico City, the damage belt observed in
Kobe and its explanation in relation to basin-edge effects
(Kawase, 1996) were among the first (dramatic) experi-
mental evidence of the engineering importance of such multi-
dimensional effects. Numerical modeling had indeed shown
for some time that the occurrence of such phenomena were
very likely due to the generation of local surface waves arising
from thickness variations of the surficial sediments which
might lead to a trapping of these waves inside the sedimen-
tary layers, resulting in larger spectral amplifications and
lengthened durations. The present study will follow an at-
tempt to go one step further to obtain quantitative informa-
tion on the actual importance of 2D or 3D site effects by
additionally accounting for the lengthening of the duration of
ground motion by considering changes in the phase of the
site-response function.

In this article, we reinvestigate site amplification in
heterogeneous media using the discrete-time correction of
ground motion at soft-soil sites affected by site effects. We
replace the traditional frequency-domain analysis by apply-
ing a set of time-domain recursive filters which, in turn,
allow the processing of data in real time, sample by sample,
as the signal comes in. For modeling site effects, a discrete
model describing the local amplification is used. In partic-
ular, we consider site-specific amplification accounting not
only for the modulus, but also for phase modifications. To do
this, we first estimate the amplitude of site response with re-
spect to a nearby hard-rock site by means of a standard spec-
tral ratio approach. In addition, we account for site-specific
modifications, which are interpreted as changes in signal
phase related to local site conditions, by making use of the
mean-group-delay algorithm proposed by Sawada (1998).
By transforming the complex standard spectral ratios to a
recursive filter, we are able to model in real time the ground
motion at soft-soil sites, thereby constraining the level and
duration of ground motion over a wide range of frequen-
cies as well as accounting for multidimensional site effects.
As shown in the following, this approach is particularly suit-
able for basins where a classical 1D numerical approach
might fail.

Calculation of the Complex Site-Response Function

The construction of the causal digital filter being used to
approximate the ground motion at the studied site is based on
an empirical complex-valued site-response function for the
frequency range of interest, meaning that an adequate site-
specific response function must be able to cope with all
prevalent filtering and amplification effects when starting
from the signal recorded earlier at a reference rock site. Such
A complex site-response function must therefore account for both amplitude changes as well as phase modifications, therefore serving to retain the statistical information about the wave phases (and therefore, about the signal duration and the time variation of amplitude and frequency).

A popular method to characterize site amplification has been the use of spectral ratios, introduced by Borcherdt (1970). Although the observed ground motion resulting from an earthquake can be described as the convolution of three contributing factors, namely a source factor $S_i(\omega)$ for an event $i$, a path factor $P_{ji}(\omega)$, and a site term $Z_j(\omega)$ for a site $j$, it is generally assumed that only the site term remains if the hypocentral distances to the sites are much larger than the distance between those sites and if the corresponding reference station is not affected by site effects, that is,

$$\frac{U_{\text{soil}}(\omega)}{U_{\text{rock}}(\omega)} = \frac{S_i(\omega) Z_{\text{soil}}(\omega) P_{\text{soil}}(\omega)}{S_i(\omega) Z_{\text{rock}}(\omega) P_{\text{rock}}(\omega)} \approx |Z_{\text{soil}}(\omega)| = A(\omega).$$

A large number of recent works on site effects have focused only on determining lateral variations in the level of amplification $A(\omega)$. However, while this approach, which is based on the ratio of the modulus of the Fourier spectra, allows the frequency-dependent amplitude modification of shaking (including 2D or 3D effects) to be retained, it does not allow a comparison of the phase characteristics of records. The ratio of the complex Fourier spectra might remove 2D and 3D effects due to the different arrival times of the seismic waves at the investigated site and the reference site when several earthquakes are analyzed and their results are averaged. However, the phase behavior is a characteristic attribute of any site resonance which is not simply 1D (e.g., Bard and Bouchon, 1985; Rothen et al., 2006). Such phase modifications can be obtained in a robust way using the mean group delay technique proposed by Sawada (1998), which can be considered as an extension of the classical spectral ratio method.

For estimating the complex site response, we follow a two-step approach in which the amplitude is estimated by the classical standard spectral ratio method, while the phase component of the spectrum is obtained using the mean group delay technique. The latter basically allows for each frequency to measure the arrival time of each phase in the seismogram by differentiating the phase spectrum $\Phi(\omega)$ obtained by Fourier transforming the signal, with respect to $\omega$. To retrieve this phase gradient, the phase spectrum has to be unwrapped first, meaning that absolute variations larger than $\pi$ are replaced by their $2\pi$ complement. The obtained quantity $T_{\varphi}(\omega) = \frac{\Phi(\omega)}{\omega}$, which has a time dimension, is defined as the group delay time. Comparing the group delay time estimated at the reference site with that calculated at the site of interest, site-specific modifications of the frequency-dependent lengthening of the signal can be identified.

However, Sawada (1998) and Beauval et al. (2003) have shown that this method for calculating the group delay time might result in very unstable results with large variations over small frequency ranges. Therefore, a smoothing operation of the phase spectrum has to be carried out before taking the numerical derivatives. Following the last-mentioned references, a frequency window function $W(\omega, \omega_b)$ can be used, taking into account the Fourier amplitude spectrum. For $W(\omega, \omega_b)$, the function proposed by Konno and Ohmachi (1998) is adopted, setting the adjustable factor of the smoothing function $b = 40$. The resulting group delay spectrum $\mu_{\varphi}(\omega)$, that is, the frequency-dependent lengthening of ground motion related to local site effects, is in turn estimated by subtracting the average group delay spectrum, calculated for each recording at the studied site, from the average group delay spectrum obtained for the corresponding recordings of the same component at the reference site.

Following Beauval et al. (2003), the phase of the seismic signal at the soft-soil site can be modified according to the observed group delay

$$\tilde{\Phi}_{\text{soil}}(\omega) = \Phi_{\text{rock}}(\omega) + \int_0^\omega \mu_{\varphi}(\omega') d\omega',$n which leads to the new site-specific complex site-response function

$$H_{\text{SSR}}(\omega) = A(\omega) e^{-i\tilde{\Phi}(\omega)}$$

with the amplitude spectrum $A(\omega)$ and the new phase $\tilde{\Phi}(\omega)$.

Modeling the Complex Site Response through the Superposition of SDOF Oscillators

In general, wave scattering and diffraction due to lateral heterogeneities and uneven surface conditions, as well as secondary impedance contrasts within the soft-soil layers, may lead to site-response functions with complicated shapes (e.g., Graves, 1993; Chávez-García et al., 1999) that cannot be properly modeled by empirical site amplification functions using a simple 1D assumption of vertically propagating waves. This means that such broadened and secondary peaks can rarely be matched theoretically when approximating the subsoil through the use of multilayered 1D models (Şafak, 1995). For modeling the frequency dependence of the empirically derived site-response function $H_{\text{SSR}}(\omega)$, the soil can, however, be approximated as a combination of damped linear oscillators with varying frequencies. This model is based on the classical Kanai–Tajimi model (Kanai, 1957; Tajimi, 1960) which defines the power spectral density of ground motion as the power spectral density of the absolute motion of a single SDOF-damped oscillator subjected to white-noise based motion. The white noise is assumed to represent the bedrock motion, and the oscillator simulates the behavior of the soil layer. Implicit in the model is the assumption that the bedrock is infinitely rigid. For such an SDOF model, the relative site amplification between the rock site and the soft-soil site $H_{\text{SSR}}(\omega)$ can be modeled by a combination of first- and second-order filters $F_k(s)$ (e.g., Oppenheim and Schafer, 1975; Hoshiba, 2013),
Ground-Motion Forecasting Using a Reference Station and Complex Site-Response Functions

\[ F_k(s) = G_0 k \Pi_{n=1}^{N} \left( \frac{\omega_{2n,k}}{\omega_{1n,k}} \right) \left( s + \omega_{1n,k} \right) \Pi_{m=1}^{M} \left( \frac{\omega_{2m,k}}{\omega_{1m,k}} \right) \left( s^2 + 2h_k \omega_{1m,k} s + \omega_{1m,k}^2 \right)^{1/2} \left( s^2 + 2h_2 \omega_{2m,k} s + \omega_{2m,k}^2 \right)^{1/2} \]  

(4)

in which \( s = i(2\pi f) = i\omega \) with frequency \( f \). \( N \) and \( M \) represent the number of first- and second-order filters. \( \omega_{jk} \) represents the angular frequencies and \( h_j \) are the damping ratios that critically determine the frequency-dependent behavior of the system.

In this approach, an iterative calculation procedure is carried out for modeling the complex site-response function \( H_{SSR}(\omega) \) shown in equation (3) through multiple \( F_k(i\omega) \). In a first step, the function \( F_k(i\omega) \) that best fits \( H_{SSR}(\omega) \) is calculated, meaning that the parameters \( G_0 k = 1 \), \( \omega_{1n,k} = 1 \), \( \omega_{2n,k} = 1 \), \( \omega_{1m,k} = 1 \), \( \omega_{2m,k} = 1 \), \( h_{1m,k} = 1 \), and \( h_{2m,k} = 1 \) in equation (4) are estimated for given values of \( N \) and \( M \) by fitting both the real and the imaginary part of \( F_k(i\omega) \) to \( H_{SSR}(\omega) \). As will be discussed later, \( N = M = 2 \) is used, because this value has been seen to allow stable results (Glaser and Baise, 2000). For the fitting algorithm, the least-squares method considering logarithmic frequency scales is applied, giving equal weights to both low and high frequencies (Ewins, 1984). The stability of the filters is ensured by allowing the parameters to vary only over physically realistic ranges of values, that is, \( 0 < \omega_j < 2\pi f _{Nyquist} \) and \( 0 < h_j < 1 \), meaning that we exclude overdamped systems. Moreover, to avoid a premature convergence of the fitting procedure into a local minimum, a probabilistic approach is applied that allows finding \( F_k(i\omega) \) near to the global optimal solution. This means that the fitting procedure is repeated several times by varying the initial random seed number.

Because each SDOF oscillator acts as a narrowband filter, the approximation of a single SDOF oscillator (i.e., \( k = 1 \)) will not be able to reproduce the frequency content of the simulated broadband site response. Therefore, for properly accounting for the frequency dependence of the site-response function \( H_{SSR}(\omega) \) over the entire frequency range of interest, the site-response function can be approximated by a linear combination of several SDOF systems calculated in an iterative procedure, with each SDOF system representing one specific natural mode (so-called modal superposition, e.g., Paz, 1997).

This means that for each iteration step, once a sufficient precision for one single SDOF oscillator \( F_k(i\omega) \) is achieved, the postfit residual for the \( k \)th iteration \( \text{Res}_k \) is calculated through \( \text{Res}_k = H_{SSR}(\omega) - F_k(i\omega), \) in which \( F_{k-1}(i\omega) = \text{Res}_k \). Accordingly, for each iteration step, a new set of the filter parameters given in equation (4) is calculated.

After a sufficient number of iterations, the final model can be computed by summing the individual SDOF oscillators, that is,

\[ F(i\omega) = \sum_k F_k(i\omega). \]  

(5)

This means that the response of the final multi-SDOF system to any given type of disturbance can be considered to be composed of all the modes represented by their corresponding SDOF systems (for a comprehensive overview, see for example, Parks and Burrus, 1987; Rao and Gupta, 1999), allowing the complex 3D signature of ground motion to be approximated by the sum of several free oscillations.

Calculation of the Recursive Filter Coefficients

To derive the parameters for the construction of an infinite impulse response filter describing the multi-SDOF oscillator model (equations 4 and 5), a bilinear transform is applied (Sklar, 2001). The transform allows a stable and discrete mapping of every point of the frequency response of the continuous-time filter to a corresponding point in the frequency response in the discrete-time domain by

\[ s = \frac{2}{T} \left( 1 - \frac{1}{1 + e^{j\omega T/2}} \right). \]  

(6)

thereby avoiding aliasing in the discrete-time domain.

To further account for nonlinearity, the frequencies \( \omega_{1n,k} \), \( \omega_{2n,k} \), \( \omega_{1m,k} \), and \( \omega_{2m,k} \) of the digital filter are prewarped, that is

\[ \omega \rightarrow \frac{2}{T} \tan \left( \frac{\omega T}{2} \right). \]  

(7)

The recursive filters, describing the convolution of ground motion recorded at a reference site for each \( F_k(i\omega) \) obtained for the \( k \)th iteration can be described in the time domain as

\[ y_k(t) = g_0 a_0 x(t) + a_1 x(t - T) - b_1 y_k(t - T) \]  

(8a)

for the first-order filters and by

\[ y_k(t) = g_0 a_0 x(t) + a_1 x(t - T) + a_2 x(t - 2T) - b_1 y_k(t - T) + b_2 y_k(t - 2T) \]  

(8b)

for the second-order filters (see Scherbaum, 1996; Hoshiba, 2013). \( x(t) \) and \( y_k(t) \) are the input (i.e., the signal recorded at the reference site) and output (the signal simulated at the site of interest) of the time series, that is, the modeled waveforms, respectively. Equations (8a) and (8b) clearly indicate that the filters are causal and recursive. The corresponding individual coefficients obtained for the \( k \)th iteration are given by

\[ g_0 = \frac{\tan \left( \frac{\omega_{2n,k} T}{2} \right)}{\tan \left( \frac{\omega_{1n,k} T}{2} \right)} \left[ \frac{1}{1 + \tan \left( \frac{\omega_{2n,k} T}{2} \right)} \right], \]

\[ a_0 = \tan \left( \frac{\omega_{1n,k} T}{2} \right) + 1, \quad a_1 = \tan \left( \frac{\omega_{1n,k} T}{2} \right) - 1, \]

\[ b_1 = \frac{\tan \left( \frac{\omega_{2n,k} T}{2} \right) - 1}{1 + \tan \left( \frac{\omega_{2n,k} T}{2} \right)}, \]  

(9)

for the first-order filters and by
\[
g_0 = \frac{\tan(\frac{\omega k T}{2})}{\tan(\frac{\omega k T}{2})} \left[ \frac{1}{1 + 2h_2 k \tan(\frac{\omega k T}{2}) + \tan^2(\frac{\omega k T}{2})} \right],
\]
\[
a_0 = 1 + 2h_1 k \tan(\frac{\omega k T}{2}) + \tan^2(\frac{\omega k T}{2}),
\]
\[
a_1 = 2 \tan^2(\frac{\omega k T}{2}) - 2,
\]
\[
a_2 = 1 - 2h_1 k \tan(\frac{\omega k T}{2}) + \tan^2(\frac{\omega k T}{2}),
\]
\[
b_1 = \frac{2 \tan^2(\frac{\omega k T}{2})}{1 + 2h_2 k \tan(\frac{\omega k T}{2}) + \tan^2(\frac{\omega k T}{2})},
\]
\[
b_2 = \frac{1 - 2h_2 k \tan(\frac{\omega k T}{2}) + \tan^2(\frac{\omega k T}{2})}{1 + 2h_2 k \tan(\frac{\omega k T}{2}) + \tan^2(\frac{\omega k T}{2})},
\]
for the second-order filters.

The final output \(y(t)\) can be obtained by the linear superposition of the individual \(y_k(t)\), that is,
\[
y(t) = \sum_k y_k(t).
\]

Figure 1 represents an overview of the individual steps carried out for the construction of the recursive filters.

**Application of the Proposed Method**

**Complex Site-Response Calculation**

In 2012, a detailed site-effect investigation was carried out in the Tajik capital Dushanbe (for details, see Pilz et al., 2013). A temporary seismic network was installed for seven months, consisting of 71 sites within the urban area covering varying geological conditions (Fig. 2, bottom). Within the city and its southern peripheral parts, the thickness of the Quaternary deposits can be up to several hundreds of meters, but it decreases to a few meters at the foothills and in areas where the bedrock outcrops. In the eastern part of the city, the soil conditions differ due to the presence of thick macropores loess-type soils having high-subsidence features under watering. Moreover, besides the network within the city limits, an additional station was installed on outcropping Paleozoic rock around 20 km north of the city which served as a reference station.

Using 184 earthquakes (3.3 \(\leq M \leq 6.7\), see Fig. 2, top) recorded during this time, the average complex-valued spectral ratios for the network sites have been derived. We only included earthquakes in the analysis with an epicentral distance to the network of at least five times the maximum interstation distance, to minimize source and path effects for waves traveling to the stations.

Figure 3 shows an example of the group delay spectrum due to site effects, that is, the frequency-resolved difference between arrival times at the soft-soil site and the reference site. Moreover, the figure displays the complex standard spectral ratios for several sites of the network obtained by considering the average group delay time. Although the modulus of the site-response function remains unchanged when applying the group-delay method, there are significant site-specific modifications in-phase which can be interpreted as changes in signal duration. As long as the source and path effects cancel out each other, such changes can be related to site geology (Beauval et al., 2003). The increase in group delay appears closely related to the local site conditions because the maximum group delay times almost systematically occur at each site’s fundamental resonance frequency. On the contrary, in general, at higher frequencies the group delay time show rather small values, meaning that the phase changes are no longer significant, even though some amplification effects might still exist.

**Calculation of the Recursive Filter Coefficients**

Once the complex site-response functions have been obtained, an iterative procedure is applied for fitting the real and the imaginary part of \(F_k(i\omega)\) (equation 4) to the corresponding empirical site-response function. As shown in Figure 4, although for the high-frequency range \(f > 10\) Hz an accurate fit can hardly be obtained, 15 iteration steps is generally sufficient for obtaining stable results for matching both the real and the imaginary parts of the site-response function up to frequencies around 10 Hz. As shown in the figure, the generation of very narrow high-peaked SDOF functions \(F_k(i\omega)\) during the iteration process can even cause the complex residuals (calculated through the vector sum of the real and the imaginary part) to increase during the iteration process (see, for example, the frequency range around 0.5 Hz for station DU03 in the left column of Fig. 4). Such high-peaked \(F_k(i\omega)\) values are usually diminished in the next iteration step by a similar mirror-like \(F_k(i\omega)\) with the opposite sign. Using a sufficient number in the order of 15–20 iterations, the complex site-response function can be fitted well by the corresponding sum of the individual SDOF oscillators, as can be seen by only a slight decrease of the complex residuals for all network sites (Fig. 4b).
Application of the Recursive Filter for the Calculation of Ground Motion

The superposition of the individual recursive filters according to equation (11) enables the direct calculation of waveforms of a soft site when the waves are initially recorded first at a reference site. However, the recursive structure given by equations (8a) and (8b) clearly indicates that the correct simulation of the different seismic phases at the soft-soil site is dictated by the arrival of corresponding phases at the reference site. This means that in such a situation, the unit sample interval at the reference site does not need to cover exactly the same time span as the unit sample interval at the soft-soil site.

Figure 2. (a) Distribution of earthquakes indicated by small dots, which were recorded by the temporary seismic network, used for the estimation of site-response functions relative to a reference station on bedrock close to the city. The main active faults are shown as thin lines. The star indicates the location of the simulated local $M_{4.1}$ event (17 May 2012). (b) Installation sites of the temporary network on different surface soil materials (Zodotarow et al., undated) within the urban area of Dushanbe. Network sites mentioned in the text are enlarged and labeled. The reference station REF was installed around 20 km north of the city. The color version of this figure is available only in the electronic edition.

Figure 3. (top) Average mean group delay for network sites (left) DU03, (middle) DU06, and (right) DU30. Thin lines show the corresponding standard deviations. (bottom) Corresponding average site-response functions for the network sites before (solid lines) and after (dashed lines) applying the group-delay method for the real (dark gray lines) and imaginary (light gray lines) parts of the site-response functions. The thick black line represents the modulus of the average site-response function, which remains unchanged. See Figure 2 for the site locations. The color version of this figure is available only in the electronic edition.
due to the variable hypocentral distances. Hence, to account for the different travel times between the hypocenter of the earthquake and the various network sites, the modeled time series has to be stretched or compressed. It may be emphasized here that the objective of stretching is not to generate new data or to account for the lengthening of ground motion at soft-soil sites, but to simply stretch the series while retaining its essential features of the different seismic phases.

Figure 4. (a) (top row) Real and (middle row) imaginary part of the single-degree-of-freedom (SDOF) oscillator of the \( k \)th iteration (thin black lines) and corresponding sum of the individual SDOF oscillators (thick black line) fitted to the complex site-response function (thick light gray line) for network sites DU03 (left column), DU06 (middle column), and DU30 (right column). The relevant moduli are shown in the bottom row. For each plot, a total of 20 SDOF functions were used. (b) Corresponding complex residuals for the \( k \)th iteration. The color version of this figure is available only in the electronic edition.
Knowing the location of an earthquake and, accordingly, the hypocentral distance to the soft-soil sites and the reference site, the stretching factor \( \tilde{s} \) is calculated in real time as the quotient of the difference in the theoretical travel time \( t \) of the \( P \) and \( S \) waves to the soft-soil and reference sites

\[
\tilde{s} = \frac{t_{S_{\text{soil}}} - t_{P_{\text{soil}}}}{t_{S_{\text{ref}}} - t_{P_{\text{ref}}}}
\]

using a simplified crustal reference model IASP91 (Kennett and Engdahl, 1991). Employing a spline function for the interpolation of the recorded time series at the reference site \( x(t) \), the time axis can be rescaled at the desired interval following Ghose (1984). Finally, the rescaled time axis for \( x(t) \) enters into the infinite response filter (equations 8a and 8b) allowing the real-time modeling of the waveforms for the soft-soil site and ensuring the later arrival of the earthquake waves at the target sites to be taken into account.

Figure 5 presents an example of the application of the recursive filter by simulating the waveforms from a local \( M_{4.1} \) earthquake that occurred on 17 May 2012 northeast of the city at a distance of 96 km to the reference site. The distance to network sites DU03, DU06, and DU30 is 106 km, 108 km, and 97 km, respectively, resulting in stretching coefficients of 1.08, 1.09, and 1.00. We also compare the recordings with the simulations of an \( M_{5.0} \) event (12 September 2012, distance to reference site 395 km). For the three network sites shown in Figure 2, the duration of 10 recorded local and regional earthquakes were divided by the corresponding simulated durations. Results of the mean duration ratios and the respective standard deviations are shown in Table 1. The table also lists the

![Figure 5](image-url)

**Figure 5.** Recordings of (top row) a local \( M_{4.1} \) event (17 May 2012) and (bottom row) a regional \( M_{5.0} \) earthquake (12 September 2012) at network sites (left column) DU03, (middle column) DU06, and (right column) DU30. Each plot shows the north–south component recording at the reference site (top waveform), the stretched waveform at the reference site (second waveform from top), the simulated waveforms for the soft-soil site (third waveform from top), and the recorded ground motion at the soft-soil site (bottom waveform).
Table 1

Mean Duration Ratios and Peak Ground Velocity Ratios

<table>
<thead>
<tr>
<th>Station</th>
<th>Duration Ratio</th>
<th>PGV Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>DU03</td>
<td>0.88 ± 0.26</td>
<td>0.86 ± 0.07</td>
</tr>
<tr>
<td>DU06</td>
<td>1.09 ± 0.16</td>
<td>0.81 ± 0.12</td>
</tr>
<tr>
<td>DU30</td>
<td>0.87 ± 0.34</td>
<td>0.91 ± 0.14</td>
</tr>
</tbody>
</table>

Mean duration ratios and mean peak ground velocity (PGV) ratios (observed divided by simulated) as well as the corresponding standard deviations using 10 local and regional earthquakes for network sites DU03, DU06, and DU30.

PGV ratios for the corresponding events as both parameters can provide confidence on how well the simulations match the recorded time series (Conte et al., 1992; Ólafsson and Sigbjörnsson, 1995).

As can be seen, the results approximate well (within uncertainty) the real observations, with a slight overestimation of the simulated events in terms of amplitude and duration for most network sites. Whereas local events occurring with a lower magnitude are accompanied by only a slight lengthening of the duration of ground motion, which can be matched very well along with relative standard deviations less than 10%, a significant increase in lengthening is observed for regional events, which is accompanied by a larger scatter around the measured values.

Discussion

A realistic simulation of ground motion relies on the exact knowledge of a large number of various parameters, from source and path effects to detailed information about the local site conditions. Small changes in any of these parameters may have a large impact on the modeled ground motion. Although pure statistical approaches can be applied, the problem of minimizing the modeling error favors models that are parameterized based on physical information.

To this regard, the concept of the standard spectral ratio method which has been introduced mainly for determining the spectral amplification at a soft-soil site relative to a nearby reference station on hard rock has been applied. The calculations of the site-response functions are based on an adequate number of earthquakes at different distance ranges (local, regional, and teleseismic for covering the entire frequency range of interest), which had been recorded at soft-soil sites as well as at the reference site, and for which the hypocentral distances were at least five times the distance between the stations of the network. Events at smaller hypocentral distances are excluded in this study to satisfy the approximations made in equation (1). Whereas the spectral ratios have been derived using the S-wave portion only, the recursive filters are applied to the whole waveforms. Although, in accordance with our tests (not shown), previous studies (e.g., Field and Jacob, 1995; Bonilla et al., 1997; Parolai et al., 2004) reported that spectral ratios of S-wave portions are similar to spectral ratios over longer time windows, we cannot exclude the possibility that the S-wave window, that is, the most energetic part of the seismogram, might reflect a slight overestimation of the simulated events in terms of amplitude and duration for most network sites (compare with the values shown in Table 1). Although the complex spectral ratios used to estimate localized site effects may be generalized using a 3 x 3 matrix transfer function between the three-component motion observed at the reference site and that at a nearby soft-soil site (Tumarkin, 1998; Paolucci, 1999), the importance of such effects for the real-time modeling of waveforms need to be better evaluated when capturing rapidly the mean wavefield features.

Once the complex site-specific spectral ratios have been obtained, the simulation of ground motion for the soft-soil sites is extremely straightforward by passing a waveform sequence, which is first recorded at a reference (i.e., hard rock) site, through recursive filters as described by the equations above. Knowing the location of the hypocenter, the simulations of ground motion are independent of the distance between the soft-soil site and the reference site as the stretching algorithm modifies the waveforms to account for the different travel times between the hypocenter of the earthquake and the various network sites. The errors are of the order of a few seconds for the P-wave onset although stretching coefficients with corrections of up to 22 s for the presented data set have been determined. Such small differences might arise from the use of a simplified crustal model that does not account for the uppermost sedimentary layers around the city of Dushanbe.

The presence of such low-velocity layers does not only amplify the level of ground motion, but it is also accompanied by a consistent increase in the mean group delay. To this regard, an extension of the classical spectral ratio method has been used (Sawada, 1998) to account for changes in the signal phase. This is particularly important for sites on top of sedimentary basins and within narrow valleys, where a classical 1D description of site amplification effects will fail. In this case, the vertical incidence of seismic waves might lead to a notable diffraction of the waves at the borders of the basin, causing the prolongation of ground motion to be significant inside the basin with a maximum value around the fundamental frequency of amplification (top panels in Fig. 3). As phase changes for the same network site were found to be similar for weak as well as for moderate events, one may consider applying this method also for strong ground motion, which is especially important in the low-frequency range.

This means that once the site parameters are known, the site-response function and the recursive filter coefficients can be determined using an appropriate number of first- and second-order filters (equation 4). For the calculations outlined above, the number of filters has been fixed to $N = M = 2$. Although numerically we are only limited by computational power as to how many modes we are able to calculate, Glaser and Baise (2000) have shown that any estimate above the second or third mode might be tenuous. Given that the data
Ground-Motion Forecasting Using a Reference Station and Complex Site-Response Functions

Moreover, as the spectral content of the first few seconds generated 3D effects due to inhomogeneities of the medium. The shortness of the recording at the reference site, which is limited in frequency (top panels in Fig. 6), significantly longer shaking over a wider frequency range is found both for the recording and the simulation at the soft-soil site, which means that local site effects influence the level of ground motion.

The time-frequency plots indicate that, in contrast to the reference site, large body-wave arrivals cause significant amplification in the intermediate frequency range between 1 Hz and 5 Hz at the soft-soil sites. For most network sites, the S-wave arrivals show a slightly dispersive behavior (clear low-frequency arrivals followed by higher frequencies) around 2 Hz both for the recordings and for the simulated waveforms, especially for the east–west component (right column in Fig. 6). Such phases are hardly observable at the rock site (Fig. 6 top). The frequency range, for which a significant level of amplification for late arrival times is present relative to the reference site, is consistent with the group delay spectrum shown in Figure 3.

This means that the discrete time-series modeling has the potential for identifying anomalous features like locally generated 3D effects due to inhomogeneities of the medium. Moreover, as the spectral content of the first few seconds after the P-wave arrival can already provide timely warning for strong-motion events close to the target city, the use of the method might contribute to improving earthquake early warning systems (Stankiewicz et al., 2015). Once the location and the back azimuth of the earthquake are known, a fast calculation of the stretching factor is possible, allowing an appropriate calculation of the delay of the arrival of the seismic waves at the soft-soil sites. Based on the real-time modeling of the waveforms at the target sites, scaling relations between the modeled P-wave amplitude and the expected peak ground shaking (e.g., Wu and Kanamori, 2005) could have the potential to further improve earthquake early warning. However, one always has to consider that the waveforms at the target sites are, in general, being reproduced well using an average site-response function, meaning that statistically one should see all individual site characteristics. However, accounting also for the azimuthal dependence of the earthquake will further ensure not missing some single event-specific features for some single events. To this regard, an adequate number of earthquakes at different distance ranges and different azimuths must be available in the used data sets.

Finally, to provide a more accurate assessment of the capability of the filter for modeling ground motion at intermediate to high frequencies, we calculate horizontal acceleration response spectra (5% damping) for the earthquake recordings and the simulated time series for the local 17 May 2012 event (Fig. 7). Response spectra portray the response of a damped SDOF oscillator to the recorded ground motion, therefore representing how a structure will react to the expected ground motions. Therefore, they are an effective measure, as they permit assessment of the influence of the various factors affecting the response and provide an approximate estimate of the response also for complex environments (Nau and Hall, 1984).

In general, both curves are in good agreement, both for shorter and longer periods. Whereas both types of spectra (compare with Fig. 3 for the Fourier spectral ratios) show similar characteristics around the fundamental frequency, the trend is different in the high-frequency range. This might be due to the response of the SDOF oscillator, causing information from the lower-frequency part of the input ground motion entering into the high-frequency part of the response spectra, implying less damping. This high-frequency behavior cannot be simulated well for all sites. As we are also capable of properly adopting the phase of earthquake ground motion, the method can provide compatible response spectra information.

Conclusions

This article demonstrates the use of parametric time-series models for forecasting the level of ground motion using the recordings at a reference site on hard rock and accounting for the effects of shallow sedimentary structures, using complex site-response functions. The mean group delay algorithm, an extension of the classical standard spectral ratio technique, allows for the calculation of modifications in the signal phase which can be interpreted as being related to local site conditions. The spatial resolution for forecasting ground motion is limited only by the number of network sites.
Figure 6. The normalized $S$ transforms of the (left column) vertical component, (middle column) north–south component, and (right column) east–west component of the M 4.1 (17 May 2012) event for the reference site recordings (top row). The plots in rows 2, 4, and 6 show the $S$ transforms of the recorded waveforms for network sites DU03, DU06, and DU30, respectively. Plots in rows 3, 5, and 7 represent the $S$ transforms of the modeled time series for network sites DU03, DU06, and DU30. The corresponding waveforms are shown in Figure 5. The color version of this figure is available only in the electronic edition.
for which the site-response functions are determined. Although the determination of the site-response functions may be time consuming, once the complex spectral ratios are obtained, an iterative procedure is applied for fitting the real and imaginary parts of multiple SDOF oscillators to the corresponding empirical site-response function, which allows the site conditions to be well mapped over a broad frequency range. The estimated filter parameters can easily be calculated by transforming the complex standard spectral ratios to an infinite response filter. The application of a recursive filter allows the forecasting of the waveforms at the soft-soil sites almost in real time, thereby constraining the level of ground motion over a wide range of frequencies. Although there might be significant travel-time differences between the various seismic phases between the hypocenter and the studied sites, their arrival time can be accurately determined by stretching or compressing the recordings at the reference site, therefore allowing the waveforms at the target site to be reproduced in terms of arrival time, amplitude, frequency content, and duration from the waveforms at a reference site. Because the duration of strong ground motion is an important parameter characterizing the total released energy, it could be shown that both the frequency content as well as the extension in duration can be accurately matched even in cases of complicated amplification patterns which might significantly increase the duration of shaking at soft-soil sites. However, although in reality the site amplification might be even more complicated and time dependent (e.g., nonlinearity phenomena) such phenomena have not been considered in the present study. In future work, the construction of an optimal recursive filter accounting for the real-time correction of site amplification factors might even enable the forecasting of ground motion if complex site conditions are present.

Data and Resources

In 2012, the Helmholtz Center Potsdam, German Research Center for Geosciences in cooperation with the Central Asian Institute for Applied Geosciences in Bishkek, Kyrgyz Republic, and the Institute of Geology, Earthquake Engineering and Seismology of the Academy of Sciences of the Republic of Tajikistan deployed a temporary seismic network covering 71 sites in the city of Dushanbe. Data recorded by this network are used in the present work.

Acknowledgments

We thank two anonymous reviewers for their constructive comments. Kevin Fleming kindly revised our English. Instruments were provided by the Geophysical Instrumental Pool Potsdam (GIPP). This research was supported by the Global Change Observatory Central Asia of the GeoForschungsZentrum (GFZ) in the framework of the Earthquake Model Central Asia (EMCA).

References


