A probabilistic approach for the construction of regional earthquake response spectra

E. Kallinikidou [a,*], S.F. Masri [a], R.L. Nigbor [a], A.W. Smyth [b], K.B. Olsen [c]

[a] Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089-2531, USA
[b] School of Engineering & Applied Science, Columbia University, New York, NY 10027, USA
[c] Department Geological Sciences, San Diego State University, San Diego, CA 92182-1010, USA

ABSTRACT

Five scenario earthquakes plausible for the Los Angeles metropolitan region, and one numerical approximation of the 1994 Northridge M_w 6.7 event, provide the database of the proposed methodology that is applied for the construction of regional earthquake response spectra. The methodology involves two main stages of data compaction. In the first stage, the Karhunen–Loève (K–L) decomposition of the excitation temporal covariance matrix is performed. In the second stage, the dominant eigenvectors are analytically approximated with Chebyshev polynomials, thus being converted from eigenvectors to eigenfunctions. This compact analytical representation of the nonstationary excitation data provides an exact closed-form solution for the nonstationary response of linear multi-degree-of-freedom systems. Furthermore, statistical inference analysis for the response variables is conducted, which leads to the construction of regional probabilistic response spectra based on the Log-Normal probability model for the response variables.

1. Introduction

A key feature of the ground motion generated by earthquakes is its transient nature. A typical accelerogram is characterized by a period of intensity growth, followed by an interval of almost steady strong shaking and finally a period of decay. For an accurate representation of a system's dynamic response to earthquake excitation, the earthquake loads should be formulated as a nonstationary stochastic process. A considerable amount of research into the simulation of nonstationary random processes was performed in the past. Representative publications in the field of applied mechanics, which deal with the stochastic response of dynamic systems, include the work of Caughey and Stumpf [1], Corotis and VanMarcke [2], Masri [3], Spanos and Lutes [4], Debchaudhury and Gasparini [5], Shinozuka and Deodatis [6], Iwan and Hou [7], Soong and Grigoriu [8], Lin and Cai [9], and Lutes and Sarkani [10].

A popular procedure for analytically representing stochastic, earthquake ground motion time histories is to model them as zero-mean, white noise processes modulated by a deterministic

envelope function representing the variation of the mean-square intensity with time. While this approach is a step in the right direction, it is not an authentic representation of actual earthquake spectral contents. On the other hand, when the earthquake ground motion is modeled as filtered white noise – where the filter may have constant or time-varying parameters – the nonstationary solutions do not have a closed form (Conte and Peng [11]; Papadimitriou and Beck [12]).

Masri and Miller [13] proposed a procedure for the compact covariance kernels' probabilistic representation (zero-mean, Gaussian probability density function modulated by a deterministic exponential function) in a form that results in the analytical evaluation for the transient, mean-square response of a linear single-degree-of-freedom (SDOF) system. This procedure was later improved by using the orthogonal Karhunen–Loève (KL) expansion and the least-squares approach to develop an approximate, analytical fit for the KL-eigenvectors of the random excitation process (Traina et al. [14]). A few years later, a relatively small ensemble of ground motion records (66) from the Los Angeles region for the 1971, San Fernando M_w 6.6 earthquake served as the application database of the method (Masri et al. [15]). Finally, high-quality ground motion records from the Los Angeles region due to the 1994, Northridge M_w 6.7 earthquake furnished an excellent excitation stochastic process for a multi-degree-of-freedom (MDOF) response solution analysis (Masri et al. [16]).

In this paper, the potential of the analytical method cited in the previous studies to provide a procedure for the construction

* Corresponding author. Tel.: +1 310 3092891.
E-mail addresses: Elena.Kallinikidou@3ds.com, ekallinikidou@solidworks.com (E. Kallinikidou), masri@usc.edu (S.F. Masri), nigbor@ucla.edu (R.L. Nigbor), smyth@civil.columbia.edu (A.W. Smyth).

of regional response spectra is addressed. Synthesized acceleration records from several scenario earthquakes plausible for the Los Angeles basin serve as the excitation stochastic database.

1.1. Motivation

In recent years, major earthquake research projects have been focusing on developing large-scale computational simulations of earthquake scenarios for geographic regions that span hundreds of kilometers. Long-term goals of such simulations include a comprehensive physics-based understanding of complex earthquake phenomena, reliable prediction of ground velocities that are expected to shake the infrastructure above ground, and identification of high-risk areas expected to sustain the worst impact. The output of such simulations yields large data files with size that reaches terabytes, making the tasks of data-management and data-archiving particularly challenging. Representative publications in the field of large-scale computational earthquake simulations include the work of Maechling et al. [17], Faerman et al. [18], Olsen et al. [19] and [20], and Benites and Olsen, [21].

This paper addresses the challenge of incorporating the emerging knowledge of this state-of-the-art earthquake research into building codes that engineers can apply to the performance-based design of earthquake-resistant structures.

1.2. Scope

As the number of available strong motion records from future earthquakes will keep increasing, due to the expansion of seismological instrumentation worldwide, the proposed compact probabilistic method can be extensively applied and its efficiency significantly enhanced. The data-processing methodology is not only a useful data-archiving and earthquake feature-extraction tool, but can also accurately quantify the average seismic risk in a probabilistic format over a spatially extended area. Eventually, the methodology can lead to the construction of regional, probabilistic design spectra that will take into account the nonstationarity of both the earthquake excitation and the response.

Synthetic accelerograms of five scenario earthquakes for the Los Angeles basin, and one numerical approximation of the 1994 Northridge M_w 6.7 event provide an excellent database to test the methodology. The extreme root-mean-square (rms) spectra for a linear SDOF system are directly calculated from the analytical solution of the nonstationary response, and compared with ensemble mean response quantities derived from the conventional response spectrum analysis. To construct probabilistic response spectra with a prescribed confidence interval, the appropriate probability model that could best describe the distribution characteristics of the response spectra variables (accelerations, velocities, and displacements) needs to be determined. A statistical inference analysis of the response parameters is performed towards that goal.

The material is organized as follows:

- Section 2: Analytical formulation for nonstationary excitation propagation. Derivation of closed-form solution for the response covariance of an MDOF system due to nonstationary support excitation.
- Section 3: Description of the Los Angeles basin model, and the earthquake scenarios' simulation data.
- Section 4: Derivation and comparison of synchronous and asynchronous earthquake acceleration covariance matrices. Orthogonal decomposition of synchronous acceleration covariances, and eigenvector approximation with Chebyshev polynomials.
- Section 5: Response spectra construction for each earthquake scenario based on (a) extreme standard deviation values, and (b) mean response values across the records' ensemble.

- Section 6: Investigation of appropriate probability model for response spectra variables. Construction of probabilistic response spectra. Superposition of extreme rms and percentile curves.
- Section 7: Conclusions and recommendation for future research.

2. Formulation

The data compaction of the system's input excitation is performed in two steps:
(1) The spectral decomposition of the input covariance matrix is calculated, and only the dominant eigenvectors are retained. (2) The dominant eigenvectors are then least-squares fitted with a series of Chebyshev orthogonal polynomials. This data compaction method permits the closed-form solution for a linear, dynamic system's nonstationary response to random excitation (Traina et al. [14]; Masri et al. [16]). Following is a brief overview of this procedure.

2.1. Eigenvector expansion of the covariance matrix

The covariance kernel $[C]$ for a system with support acceleration $\ddot{s}(t)$ is a symmetric, square matrix with values defined as

$$C_{\ddot{s}\ddot{s}}(t_1, t_2) = E[(\ddot{s}(t_1) - \mu_{\ddot{s}}(t_1))(\ddot{s}(t_2) - \mu_{\ddot{s}}(t_2))] \qquad (1)$$

where $E[\cdot]$ is the expectation operator, and $\mu_{\ddot{s}}(t)$ is the mean value of the support acceleration at time t. Using the Karhunen–Loève expansion, the spectral representation of the $[C]$ matrix of order n by n may be expressed as

$$[C] = \sum_{i=1}^{n} \lambda_i \mathbf{p}_i \mathbf{p}_i^{\mathrm{T}} \qquad (2)$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ are the real and positive eigenvalues of $[C]$, and $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are the corresponding normalized eigenvectors such that $\mathbf{p}_i^{\mathrm{T}} \mathbf{p}_j = \delta_{ij}$, for $i = 1, 2, \ldots, n$. In most cases, only a small fraction k out of a total n eigenvalues are likely to reconstruct the covariance matrix $[C_k]$. Matrix $[C_k]$ represents the reconstruction of the covariance matrix by using the first largest eigenvalues k out of a total n eigenvectors, and matrix $[E_k]$ contains the error due to the eigenvector truncation:

$$[C] = \sum_{i=1}^{k} \lambda_i \mathbf{p}_i \mathbf{p}_i^{\mathrm{T}} + \sum_{i=k+1}^{n} \lambda_i \mathbf{p}_i \mathbf{p}_i^{\mathrm{T}} = [C_k] + [E_k]. \qquad (3)$$

2.2. Least-squares fit of eigenvectors

The second step in the data condensation procedure involves the approximation of the truncated set of k discrete eigenvectors with continuous analytical functions. Chebyshev polynomials were chosen for the least-squares approximation because, in addition to being orthogonal, they have the desirable feature of equal-error approximation within an interval of interest, as well as being convenient for defining analytical solutions for the dynamic system response.

The approximated covariance matrix $[\hat{C}_k]$ can then be reconstructed, and expressed in terms of a Chebyshev polynomial series (Masri et al. [15]):

$$[\hat{C}_k(t_1, t_2)] = \sum_{i=1}^{k} \lambda_i \sum_{j=0}^{m_i-1} \sum_{\ell=0}^{m_i-1} H_{ij} H_{i\ell} T_j(t_1') T_\ell(t_2'), \qquad (4)$$

where the T's are the Chebyshev polynomials, H_{ij} is the coefficient of the Chebyshev polynomial of order j associated with the eigenvector $\mathbf{p}_i$, m_i is the number of Chebyshev polynomials used, and $0 \leq t_i \leq t_{\max}$ for $i = 1, 2$.

2.3. Nonstationary excitation propagation

The analytical representation of the input acceleration covariance matrix (Eq. (4)) permits the closed-form response solution for a multi-degree-of-freedom (MDOF), linear, viscously damped (and not necessarily of the proportional type), dynamic system. The system's response is arranged in a $2n$-dimensional vector of the form

$$\{y(t)\} = \left\{ \begin{array}{c} x(t) \\ -\,-\,- \\ \dot{x}(t) \end{array} \right\} \tag{5}$$

where $\{y(t)\}$ is the state vector, and defines the state of the system uniquely for any time t. In this application, the n-dimensional vectors $x(t)$ and $\dot{x}(t)$ are the nodal, relative displacements and velocities with respect to the support position. The system of differential equations of motion for an MDOF linear system excited by a base excitation $\ddot{s}(t)$ is given in matrix form by

$$[M]\{\ddot{x}(t)\} + [C]\{\dot{x}(t)\} + [K]\{x(t)\} = -[M]\{e\}\ddot{s}(t) \tag{6}$$

where $[M]$, $[C]$, and $[K]$ are the $n \times n$ mass (diagonal), damping, and stiffness matrices, respectively, and $\{e\}$ is the n-dimensional unit vector.

To derive the closed-form solution for the response covariance matrix of a MDOF system due to a support nonstationary excitation, analytical representation of Chebyshev polynomials and integrals in a finite series form are constructed. Since the focus of this paper is the application of existing analytical tools to develop seismic response spectra, the intermediate steps of the theoretical derivation are omitted. A detailed treatment of the closed-form solution of the nonstationary response can be found in the thesis by Smyth [22] and in the work of Masri et al. [16]. The development of the seismic response spectra in this paper is based on an SDOF representation of the response covariance.

3. Application

3.1. Los Angeles basin model and earthquake scenarios

The metropolitan area of the greater Los Angeles region lies on top of a deep sedimentary basin. During past earthquakes, such as the 1985, M_w 8.1 Michoacan, and the 1989, M_w 6.9 Loma Prieta events, it was observed that ground motions in basins get significantly amplified, as was the case in Mexico City, and the Marina district of San Francisco. To obtain a confident estimate of a particular site's basin response, all the possible regional earthquake scenarios need to be included in a probabilistic seismic hazard analysis. In this study, the following five scenario earthquakes, all plausible for the Los Angeles region, and a numerical approximation of the 1994, M_w 6.7 Northridge event are examined:

(1) A blind thrust on the Elysian Park fault (EP); (2) Thrust on the Santa Monica fault (SM); (3) Northwest propagating rupture on the Newport–Inglewood fault (NI); (4) Southeast propagating rupture (SAFSE); and (5) Northwest propagating rupture on a 170 km long stretch of the San Andreas fault (SAFNW); (6) Approximation of the 1994, M_w 6.7 Northridge (NOR) earthquake.

These earthquake simulations were developed by Olsen [23] to explore the ground motion amplification of seismic waves in the deep Los Angeles basin. They are all within the "geological reasonable scenarios" defined by Dolan et al. [24]. The rupture parameters for the six earthquake simulations are given in Table 1.[1]

Fig. 1. Topographic map of Southern California. The larger rectangular area depicts the extent of the model for the San Andreas Fault (SAF) scenarios. The smaller rectangle shows the region used for modeling the remaining scenarios.
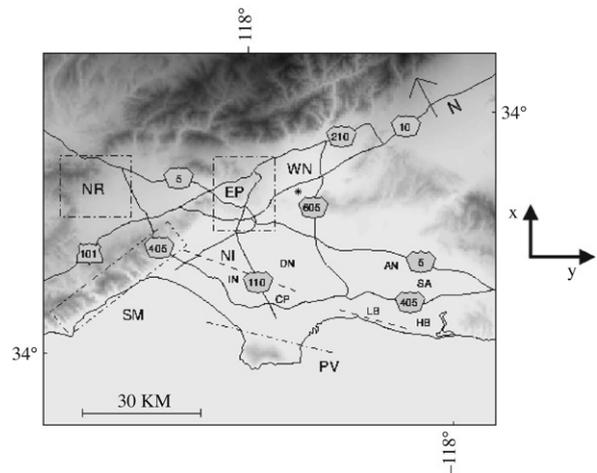
Fig. 2. Topographic map of the Los Angeles basin area with the surface projections of the fault planes associated with the earthquake scenarios: Santa Monica (SM), Elysian Park (EP), Newport–Inglewood (NI), Palos Verdes (PV), 1994 **M** 6.7 Northridge (NR).

A topographic map of southern California is illustrated in Fig. 1. The larger rectangular area depicts the extent of the model for the San Andreas Fault (SAF) scenarios. Within the smaller rectangular area, which is shown in Fig. 2 along with the surface projections of the fault planes, lies the region for which the remaining scenarios are simulated. The reference frame of the numerical model was rotated 28° counterclockwise from the North, to minimize the computational requirements for the simulations. Throughout the present study, the horizontal ground motions and the subsequent response parameters are analyzed along the azimuths of 118° (x-axis), and 28° (y-axis). The model of the extended Los Angeles basin spreads over an area of 155 km × 134 km × 34 km, and is discretized with a grid spacing of approximately 0.4 km for a total of 388 × 337 × 85 (= 11,114,260) grid points. It includes the Los Angeles basin, the smaller San Fernando basin to the northwest separated from the Los Angeles basin by the Santa Monica mountains, and the Ventura basin continuing to the North. A relatively small basin is located below the San Gabriel Valley, and the larger, but shallower

---

[1] A comprehensive treatment of the underlying physics of the wave propagation in the Los Angeles basin for the earthquake simulations can be found in Olsen [25]).
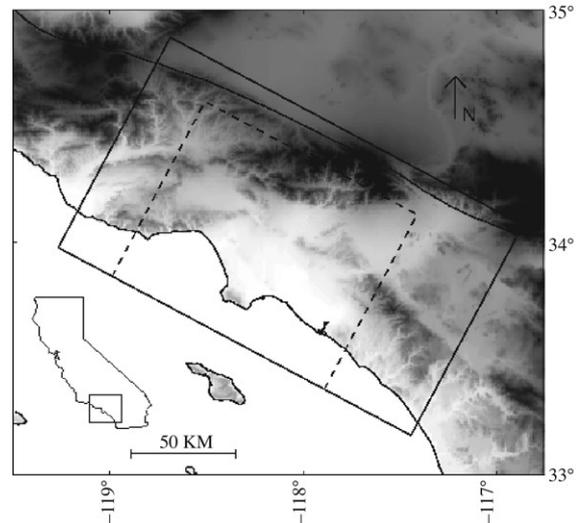
Table 1
Earthquake rupture parameters (after Olsen).

	EP	SM	NI	SAFNW	SAFSE	NOR
Hypocenter						
Longitude (deg)	−118.20	−118.64	−118.24	−118.66	−117.09	−118.53
Latitude (deg)	34.10	34.06	33.88	34.79	34.09	34.20
Depth (km)	14.8	10.4	7.6	4.0	15.2	17.6
Fault parameters						
Width (km)	21	17	15	16	16	24
Length (km)	16	37	41	170	170	18
Depth to top (km)	10.5	0.4	0.4	0.4	0.4	5.0
Dip (deg)	25	65	90	90	90	40
Strike (deg)	298	260	138	118	118	118
Rake (deg)	90	90	180	180	180	101
Moment magnitude	6.75	6.75	6.75	7.5	7.5	6.7

Table 2
Los Angeles basin 3D modeling parameters (after Olsen).

Spatial discretization (km)	0.4
Temporal discretization (s)	0.25
Lowest P-wave velocity (km/s)	2.41
Lowest S-wave velocity (km/s)	1.0
Lowest density (kg/m^3)	2070
Number of time steps	360,500 (Northridge)
Total simulation time (s)	90,125 (Northridge)

Chino basin, which extends the LA basin to the East. The topography above sea level was not included. The three-dimensional (3D) modeling parameters are listed in Table 2.

3.2. Limitations of the Los Angeles basin amplification study

According to Olsen [25], the main limitations of the average amplification map for the Los Angeles basin area were primarily due to the relatively small number of scenario earthquakes (nine in the original study) that were used in the research study and the maximum frequency (0.5 Hz) used for the ground motion simulation, and secondarily, to uncertainties concerning the accuracy of the basin model, and the omission of surface layers with shear velocities less than 1 km/s. It is important to incorporate more simulations of geologically reasonable scenarios, in order to reduce the uncertainty of the average amplification ratios,[2] and also to increase the highest frequencies, to capture the entire spectrum of amplification effects. Considering the recent explosive increase in computational power, it may be feasible to construct more accurate amplification maps in future studies with broader frequency content derived from smaller model grid spacing. These maps should include the effects of realistic rupture parameters, and anelastic attenuation as well.

3.3. Velocity records

The synthetic velocity seismograms were low-pass filtered to frequencies less than 0.5 Hz (Butterworth filter with four poles and two passes). There were three sets of records – 118° (X-direction), 28° (Y-direction), and vertical (Z-direction) – for each simulation, which brings the total number of processed sets to 18. Each velocity record corresponds to a grid point of the basin's modeled region that covers a total area of 115 km × 95 km. The total number of grid points on the ground surface is 301 × 216, resulting in 195,048 particle velocity surface records for each earthquake

scenario simulation. Each record has a 90 s duration (125 s for the Northridge records), and a sampling rate dt = 0.25 s.

The maximum absolute horizontal particle velocity magnitudes range for the horizontal directions from 0.767 m/s (San Andreas NW) to 4.168 m/s (Newport–Inglewood) and for the vertical from 0.286 m/s (San Andreas NW) to 4.278 m/s (Santa Monica). Since the narrow-band frequency content of the simulated time histories is limited to frequencies less than 0.5 Hz, and the near-surface S-wave velocity's upper bound is 1.0 km/s, these values are estimated to be lower bounds of the expected ground motion velocities in the near source area. Future simulations that would result from signals with a wider frequency content, and more realistic values for the surface S-wave velocity, would give rise to higher ground velocities and accelerations.

3.4. Acceleration records

To derive the ground acceleration time histories, a numerical differentiation by fitting a moving three-point quadratic polynomial ($y = a_0 + a_1x + a_2x^2$) to the ground velocity record data was performed. A three-point analytical fit was chosen instead of a five-point fit ($y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$), which, intuitively, would have resulted in higher accuracy, because the accuracy of the three-point fit proved to be adequate when applied to the low frequency velocity signals (4 samples/s).

The cumulative absolute velocity (CAV), a severity indicator of an acceleration time history, was calculated for each record. The CAV, defined as the area under the absolute acceleration versus time duration curve, is given by CAV = $\int_0^{t_{max}} |\alpha(t)| \, dt$. The three acceleration records with the maximum root square of the sum of the squares (RSSS) CAV ($\sqrt{CAV_X^2 + CAV_Y^2 + CAV_Z^2}$) magnitude of 40.34 m/s resulted from the Newport–Inglewood dataset. The X-direction acceleration record and the corresponding fast Fourier transform (FFT) diagram are given in Fig. 3. The narrow-band nature of the signal (< 0.5 Hz) is evident in the FFT plot. Particle acceleration propagation snapshots of the San Andreas, SE propagating rupture (X-direction) are shown in Fig. 4.

4. Synchronous and asynchronous covariances

The temporal covariance matrices of order 360 × 360 (500 × 500, for the Northridge sets) extending over a time span of 90 s (125 s, for the Northridge sets) were constructed from the velocity and acceleration record ensembles. In general, the order of the covariance matrices is not constant, but depends on the number of samples in the signal. The covariance matrices were calculated twice: first, from the unprocessed, unsynchronized records, and second, from records that were synchronized at the trigger threshold value of 1% rms for each ensemble.
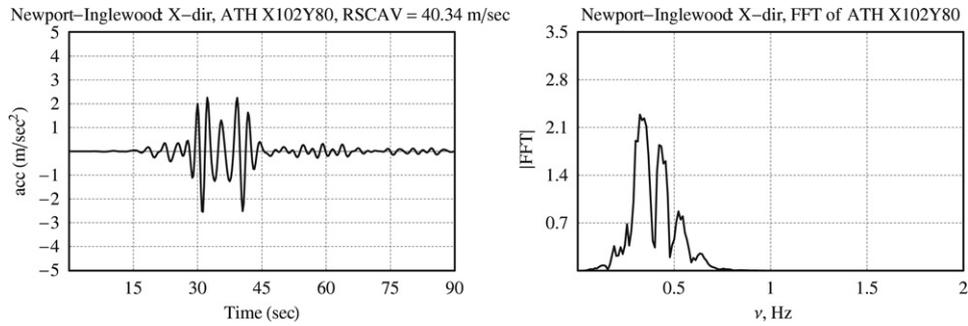
---

[2] Amplification is quantified as the peak velocity obtained from the 3D simulation divided by that predicted using a regional one-dimensional (1D) crustal model.

Fig. 3. Acceleration time history (left side) and its FFT diagram (right side) of the record with the largest root square of the sum of the squares (RSSS) cumulative absolute velocity (CAV) value. The record is from the Newport–Inglewood X-direction dataset.
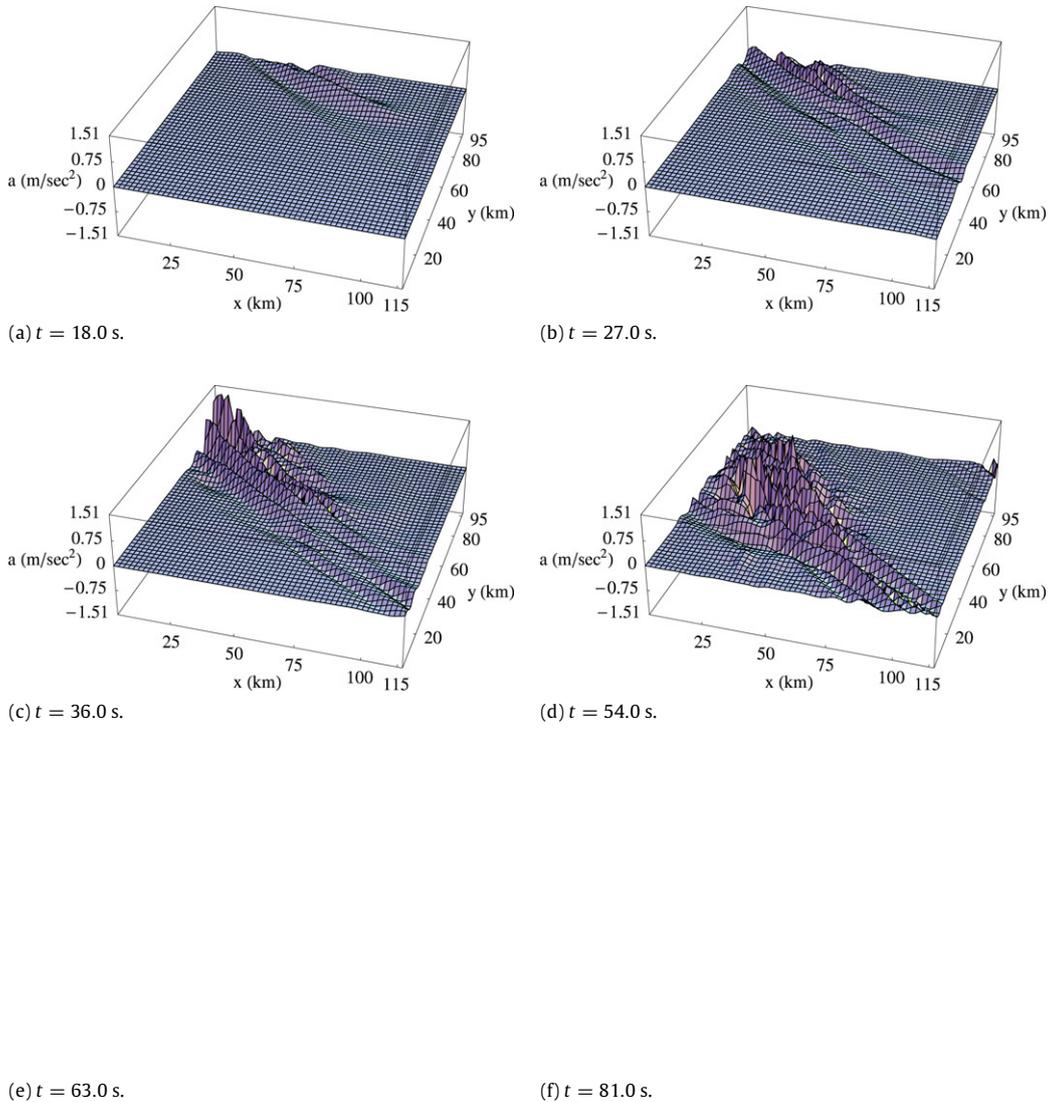

(a) $t = 18.0$ s.

(b) $t = 27.0$ s.

(c) $t = 36.0$ s.

(d) $t = 54.0$ s.

(e) $t = 63.0$ s.

(f) $t = 81.0$ s.

Fig. 4. Six 3D particle acceleration snapshots from the San Andreas, SE propagating rupture (X-direction). The horizontal axes are the basin's kilometric distances and the vertical axis is the acceleration amplitude.

The velocity ground motion records were asynchronous because they captured the wave propagation effect through the extended Los Angeles region. The synchronization of records was performed by shifting the velocities and accelerations in the time domain. The rms value for each record ensemble was calculated, and the first sample ($t = 0$) of each record was set to 1% of the ensemble rms. It was expected the record synchronization would result in a conservative estimate for the covariance amplitudes.

The synchronization of the ground motion records, indeed, gave rise to slightly larger covariance amplitudes as expected. This is clearly evident in Fig. 5(a–f), where the two variance time histories which depict the time variations of the mean-square accelerations (combined X, Y, and Z direction) are plotted for all six simulations. The covariance amplitudes from the synchronized records, besides being amplified, were also time-shifted in the case of the San Andreas NW (Fig. 5a), and San Andreas SE (Fig. 5b) scenarios.