DEVELOPMENT OF A DATA-BASED PROCEDURE FOR CONSTRUCTING REGIONAL PROBABILISTIC EARTHQUAKE RESPONSE SPECTRA FOR STRUCTURAL CONTROL APPLICATIONS

E. Kallinikidou[1], S. F. Masri[1], R. Nigbor[1], A.W. Smyth[2] and K. B. Olsen[3]

[1] Viterbi School of Engineering, University of Southern California, Los Angeles, California, USA
[2] School of Engineering & Applied Science, Columbia University, New York, New York, USA
[3] Dept. Geological Sciences, San Diego State University, San Diego, California, USA
E-mail: *kallinik@usc.edu*

ABSTRACT

Five scenario earthquakes plausible for the Los Angeles metropolitan region, and one numerical approximation of the 1994 Northridge Mw 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 MDOF 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.

Keywords: regional earthquake spectra, probabilistic earthquake spectra, Karhunen-Loéve expansion, regional seismic risk, proper orthogonal decomposition.

1 INTRODUCTION

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. 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.

Synthetic accelerograms of five scenario earthquakes for the Los Angeles basin, and one numerical approximation of the 1994 Northridge Mw 6.7 event provide an excellent database to test the proposed 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. A detailed treatment of the closed-form solution for the response covariance matrix of an n-dof system due to a support nonstationary excitation can be found in the thesis by Smyth, 1998 (7) and in the work of Masri *et al.*, 1998 (3). 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.

2 LOS ANGELES BASIN EARTHQUAKE SCENARIOS

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, Mw 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, Mw 6.7 Northridge (NOR) earthquake.

These earthquake simulations were developed by Olsen (Olsen, 1994) (5) to explore the ground motion amplification of seismic waves in the deep Los Angeles basin.[1] 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 x 134 Km x 34 Km, and is discretized with a grid spacing of approximately 0.4 Km.

3 DATA PROCESSING APPROACH

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.*, 1986 (8); Masri *et al.*, 1998 (3)).

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

[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, 2000 (6).

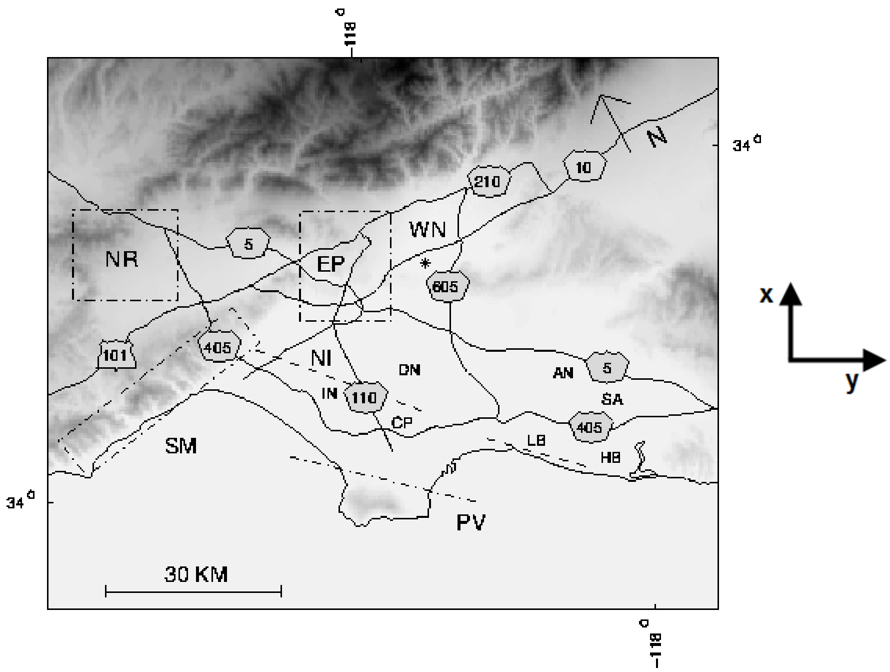

Figure 1: Topographic map of the Los Angeles basin area with the surface projections of the fault planes associated with the earthquake scenarios: Santa Monica (SM), Elysian Park (EP), Newport-Inglewood (NI), Palos Verdes (PV), 1994 **M** 6.7 Northridge (NR).

with values defined as:

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

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 in terms of a truncated set of eigenvectors, and subsequently, through the application of formal random vibration approaches, the transient mean-square system response $E[y^2(t)]$ can be analytically determined.

4 CONSTRUCTION OF RESPONSE SPECTRA

The transient mean-square response time histories of a single-degree-of-freedom (SDOF) system with a critical damping ratio of $\zeta = 0.01, 0.05,$ and 0.10 were obtained from the closed-form analytical expression for the response covariance. The six scenario earthquake approximating covariances, used as input, were reconstructed with the first 40 eigenvectors of the original covariances that had been least-squares fitted by Chebyshev polynomials of order $m = 175$. A sample mean-square time history of a representative SDOF with a natural frequency of 0.2 Hz and critical damping ratio of $\zeta = 0.01$ is shown in Fig. 2.

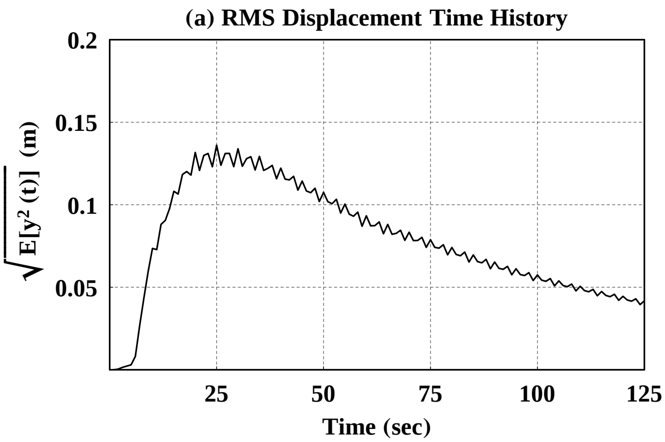

Figure 2: Transient root-mean-square response of a SDOF system with frequency of 0.2 Hz, and critical damping ratio of 1%, for the Northridge (X-dir) scenario earthquake.

The extreme values of the rms response of a damped SDOF system are given by

$$R_d\,(T, \zeta) \equiv \sqrt{\underset{t}{max}\,\{E[y^2(t)]\}} \tag{2}$$

$$R_v\,(T, \zeta) \equiv \sqrt{\underset{t}{max}\,\{E[\dot{y}^2(t)]\}} \tag{3}$$

where $E[y^2(t)]$ and $E[\dot{y}^2(t)]$ are the mean-square displacement, and velocity response time histories, respectively.

For each scenario earthquake, the $E[y^2(t)]$ and $E[\dot{y}^2(t)]$ response time histories were calculated with input acceleration covariances derived from the horizontal (combined X and Y direction), and vertical (Z direction) data sets. Representative plots for the Newport-Inglewood scenarios are shown on the left side column in Fig. 3. The amplitudes of the response values for all spectral plots are significant for periods larger that 2 sec. This result is expected due to the limited frequency bandwidth ($n < 0.5\,Hz$) of the input acceleration records.

Furthermore, for each synthetic acceleration record, the response of a SDOF linear system subjected to input acceleration was calculated by solving analytically the classic differential equation of motion in a computationally efficient fashion:[2]

$$\ddot{x} + 2\zeta\omega\dot{x} + \omega^2 x = f(t) \tag{4}$$

The maximum absolute values of the response time histories (SA: absolute acceleration, SV: relative velocity, and SD: relative displacement) were plotted versus period values T, and damping ratios $\zeta = 0.01, 0.05,$ and 0.10 to form the response spectra. Representative mean response spectra plots for the

[2]The analytical solution of the second order differential equation is given in Nigam and Jennings, 1969 (4).

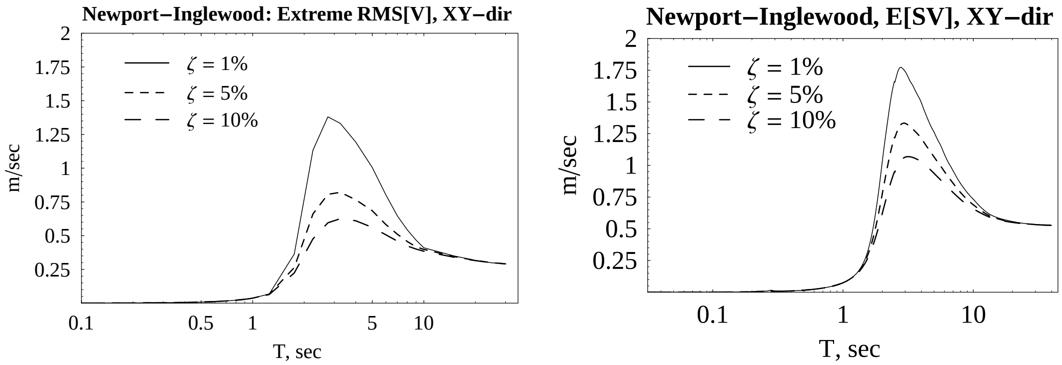

Figure 3: Comparison of extreme RMS[V] (left side), and ensemble mean spectral velocity E[SV] (right side) for the combined X and Y horizontal datasets from the Newport-Inglewood scenario earthquake.

Newport-Inglewood scenarios are shown on the right side column in Fig. 3. For ease of comparison, identical amplitude and period scales are used for the two types of spectra plots [3]. It is worth keeping in mind that the two parameters are qualitatively different: one tracks the extreme value of the response rms (obtained at one specific time instant), while the mean response spectra calculations provide the average of the ensemble response peaks, each one occurring at a different time.

5 STATISTICAL INFERENCE OF THE RESPONSE SPECTRA VARIABLES

In order to construct probabilistic response spectra with a prescribed confidence interval, frequency and cumulative distribution diagrams of the response spectral variables for $k = 1200$ intervals were constructed for all data samples. The diagrams of a linear oscillator having a natural period of 4 seconds ($\zeta = 0.01$) are presented in Fig. 4. Furthermore, the observed response sample data were plotted versus the linear probability scales of the following assumed distributions: Log-Normal, Wiebull, Frechet, and Gamma. The data points should follow a linear trend, if the assumed distribution is an appropriate model for the data. Based on the quality of the linearity that these probability plots exhibited, the following distributions were chosen as the most promising candidates for a potential suitable probability model for the spectral response variables of interest: (a) Two–parameter Log-Normal; (b) Three–parameter Log-Normal; (c) Two–parameter Weibull; (d) Two–parameter Gamma.

[3] The mean and root-mean-square response spectra plots for the remaining scenarios are available, but are not included in this paper due to size constraints.

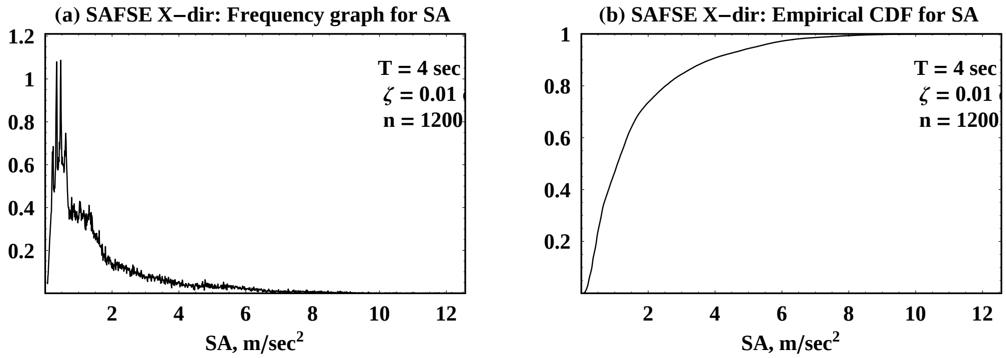

Figure 4: Frequency and cumulative distribution frequency (CDF) diagrams for the spectral acceleration of a linear oscillator with a period of 4 sec and damping $\zeta = 0.01$ for the San Andreas SE X–dir dataset).

5.1 Parameter Estimation for the Candidate Distributions

The optimal parameter estimators of the chosen distributions were determined with the *maximum likelihood* (ML) method (Ang and Tang, (1)).The parameters of the distributions were chosen from a parametric optimization process which minimized the *chi-square* quantity

$$\chi = \sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i} \tag{5}$$

where $n_1, n_2, \cdots, n_k \, (k = 1200)$ were the observed frequencies, and $e_1, e_2, \cdots, e_k \, (k = 1200)$ the assumed theoretical frequencies. The optimization procedure used in this task was based on a trial-and-error process. The interval limits were chosen through an iterative process, which predetermined the best plausible range of values for minimizing the quantity χ. A detailed discussion on the parameter estimation can be found in the thesis by Kallinikidou, 2006 (2). For all response samples, the smallest values of the error quantities were for the period values between 2 and 4 sec. By comparing the minimum error quantities, among the assumed candidate distributions—Log-Normal, Weibull, and Gamma—the Log-Normal was found to be the most suitable for the observed frequencies of all three response samples: spectral accelerations, velocities, and displacements.

5.2 Probabilistic Response Spectra Based on the Log-Normal Distribution

The two-parameter Log-Normal model yielded the minimum χ values over the period spectrum for both earthquake ensembles (San Andreas SE and Newport-Inglewood), and was selected to formulate the probabilistic response spectra. Furthermore, the 80%, 90%, and 95% percentile curves of the response variables (SA, SV, SD) based on the families of Log-Normal distributions were plotted for each record ensemble. (Fig. 5).

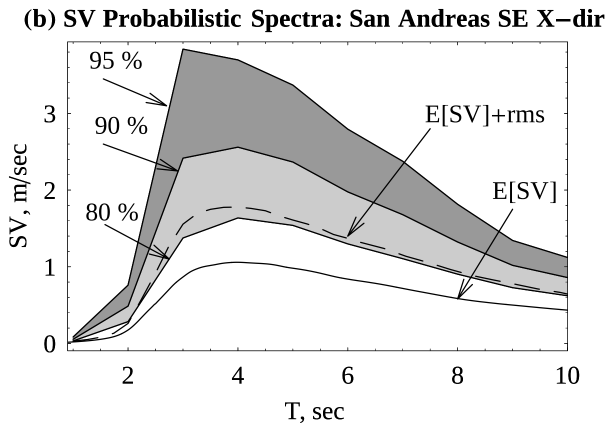

Figure 5: Response curves of spectral velocities SV over period values for the San Andreas SE X–dir dataset. The curves correspond to ensemble mean values E[SV] and percentiles: 80%, 90%, 95% of non-exceedance. The dashed lines are the ensemble mean plus rms values for the spectral velocities SV (damping $\zeta = 0.01$).

5.3 Probabilistic Association of RMS Response Measures

To inquire about a probabilistic association between the extreme rms response quantities derived from the analytical solution of the nonstationary response and the response spectral quantities derived from the statistical inference analysis, curves of multiple rms values ($k{*}rms, k = 1, 2, 3, 4, 5$) were superimposed on the response spectra percentile graphs that are based on the Log-normal distribution (Fig. 6). The rms values are the extreme values of the transient mean-square responses, and are determined from Eqs. 2 and 3 (see Fig. 2 for a sample transient mean-square response).

These graphs were constructed only for the spectral velocities SV and spectral displacements SD responses of the San Andreas SE X–dir dataset, and for one level of damping $\zeta = 0.01$). These graphs indicate the expected level of probability confidence, when the rms value (or a multiple value of rms) is used for a performance-based design of earthquake-resistant structures. For example, based on Fig. 6, by selecting the value of 5*rms (i.e., five times the extreme rms response) as the design threshold for spectral velocities SV, there is a 95% probability that this value will not be exceeded.

6 CONCLUSIONS

To construct regional earthquake response spectra from thousands of synthesized acceleration records derived from several scenario earthquakes plausible for the Los Angeles basin, a procedure with two stages of analysis is proposed. The first stage of analysis includes the construction of the excitation temporal covariance matrices, their Karhunen-Loève (K-L) decomposition, and subsequent conversion

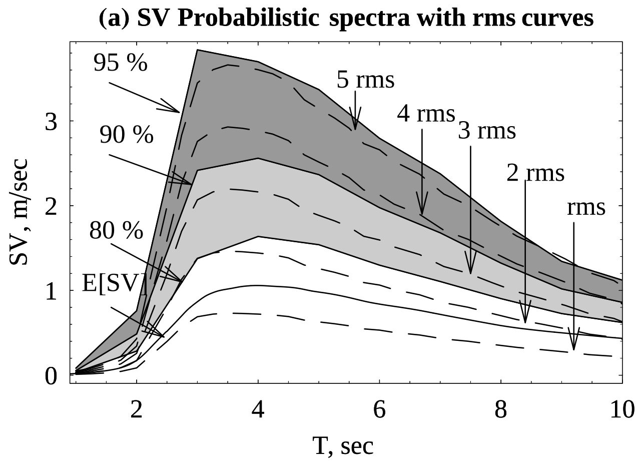

Figure 6: Superposition of percentile curves (solid lines) and extreme rms (dashed lines) for the spectral velocities SV of the San Andreas SE X–dir scenario (damping is $\zeta = 0.01$). The solid lines correspond to ensemble spectral means (E[SV]), and 80%, 90%, and 95% values of non-exceedance. The dashed lines correspond to the extreme $k * rms$ values for $k = 1, 2, 3, 4, 5$.

of the dominant K-L eigenvectors to eigenfunctions with Chebyshev polynomials. Based on the exact closed-form solution for the nonstationary response of linear MDOF systems, response spectra for different damping values are constructed for the extreme values of the rms response quantities.

For the second stage of analysis, the response of a SDOF system subjected to acceleration excitation is calculated by solving analytically the differential equation of motion in a computationally efficient fashion for processing large size data-sets. Response spectra for different damping values are constructed for the mean values of the ensemble response peaks. In addition, statistical inference analysis for the response variables leads to the construction of response spectra associated with percentile values of non-exceedance based on the Log-Normal probability model for the response variables.

A significant contribution of this study is the establishment of a quantitative relationship between the analytically-determined extreme values of the nonstationary system response, and the probabilistic response spectra directly constructed from the ensemble statistics of the available earthquake records, so as to estimate the confidence levels of non-exceedance of specific spectral response levels. Such information can be useful in conducting performance-based designs matching certain reliability constraints.

The main limitations for the application of the proposed methodology for the construction of regional earthquake spectra are primarily due to the relatively small number of scenario earthquakes considered in this 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 1.0 Km/sec.

7 ACKNOWLEDGMENT

This study was supported in part by grants from the U.S. National Science Foundation.

References

[1] A. Ang and W. Tang. *"Probability Concepts in Engineering Planning and Design"*. John Wiley & Sons, Inc., 1975.

[2] E. Kallinikidou. *Analytical and Experimental Methods for Regional Earthquake Spectra and Structural Health Monitoring Applications*. PhD thesis, University of Southern California, Los Angeles, CA, December 2006.

[3] S. F. Masri, A. W. Smyth, and M. I. Traina. Probabilistic representation and transmission of nonstationary processes in multi-degree-of-freedom systems. *ASME Journal of Applied Mechanics*, 65:398–409, 1998.

[4] N. C. Nigam and P. C. Jennings. Calculation of response spectra from strong-motion earthquake records. *Bulletin of the Seismological Society of America*, 59(2):909–922, 1969.

[5] K. B. Olsen. *Simulation of three-dimensional wave propagation in the Salt Lake Basin*. PhD thesis, University of Utah, Salt Lake City, Utah, 1994.

[6] K. B. Olsen. Site amplification in the los angeles basin from three-dimensional modeling of ground motion. *Bull. Seism. Soc. Am.*, 90(6B):S77–S94, 2000.

[7] A.W. Smyth. *Experimental and Analytical Studies in Nonlinear System Identification and Modeling for Structural Control*. PhD thesis, University of Southern California, Los Angeles, CA, May 1998.

[8] M. I. Traina, R. K. Miller, and S. F. Masri. Orthogonal decomposition and transmission of nonstationary random processes. *Probabilistic Engineering Mechanics*, 1(3):136–149, 1986.