Record Selection for Nonlinear Seismic Analysis of Structures

Iunio Iervolino[a),b)] **and C. Allin Cornell,**[b)] M.EERI

This study addresses the question of selection and amplitude scaling of accelerograms for predicting the nonlinear seismic response of structures. Despite the current practices of record selection according to a specific magnitude-distance scenario and scaling to a common level, neither aspect of this process has received significant research attention to ascertain the benefits or effects of these practices on the conclusions. This paper hypothesizes that neither these usual principal seismological characteristics nor scaling of records matters to the nonlinear response of structures. It then investigates under what conditions this hypothesis may not be sustainable. Two classes of records sets are compared in several case studies: one class is carefully chosen to represent a specific magnitude and distance scenario, the other is chosen randomly from a large catalog. Results of time-history analyses are formally compared by a simple statistical hypothesis test to assess the difference, if any, between nonlinear demands of the two classes of records. The effect of the degree of scaling (by first-mode spectral acceleration level) is investigated in the same way. Results here show (1) little evidence to support the need for a careful site-specific process of record selection by magnitude and distance, and (2) that concern over scenario-to-scenario record scaling, at least within the limits tested, may not be justified. [DOI: 10.1193/1.1990199]

INTRODUCTION

This paper presents a determinedly transparent study of the question of selection and scaling of accelerograms for predicting the nonlinear dynamic response of a structure at a specific site. The preferred current practice is to carefully select records that reflect the expected magnitude, distance, and other characteristics of the source of the events that are in some sense most likely to threaten the structure. The records are then typically scaled to some common representative level. Neither aspect of this process, neither selection nor scaling, has received significant research attention to ascertain their effects on the conclusions. This paper approaches these subjects inversely; it hypothesizes that neither the usual principal seismological characteristics nor scaling of records matters to the nonlinear response of structures. It then investigates under what conditions this hypothesis may not be sustainable. The study deals with *ordinary records*; softer soil site and specific near-fault effects, such as directivity-induced pulses, both of which may cause narrow-band response spectra are carefully avoided. Nonlinear analysis case stud-

a) Dept. of Structural Analysis and Design, University of Naples Federico II, 21 via Claudio, Naples 80125, Italy.
b) Dept. of Civil and Environmental Engineering, Terman Engineering Center, Stanford University, CA 94305

Earthquake Spectra, Volume 21, No. 3, pages 685–713, August 2005; © 2005, Earthquake Engineering Research Institute

ies consider different periods, force-displacement characteristic relationships (backbones), ductility levels, and structural types. Sets of two classes of records are compared in each case: one class is carefully chosen to represent a specific magnitude and distance scenario (a "target set"), and another class is chosen randomly from a large catalog (an "arbitrary set") and scaled to match the target set in general amplitude. Results of time-history analyses are formally compared by a simple statistical hypothesis test to assess the difference, if any, between nonlinear demands of the two classes of records. The effect of the degree of scaling (by first-mode spectral acceleration level) is investigated in the same way. Results here show (1) little evidence to support the need for careful site-specific process of record selection by magnitude and distance, and (2) that concern over scenario-to-scenario record scaling, at least within the limits tested, may not be justified. This study does not explain the role of systematic spectral shape deviations, such as those due to soft soil, directivity, or scenarios calling for non-median ground motions.

MOTIVATION AND FRAMEWORK

The study is aimed at improving the bases for guidelines for earthquake engineering practice in terms of (1) characteristic that should be taken into account in accelerograms selection, (2) scaling of records in order to get scenario (target) intensity, and (3) sufficient size of record sets. Moreover, it will also shed light on other issues such as structural period and/or backbone sensitivity.

The current state of best practice (e.g., U.S. Nuclear Regulatory Commission 2001) in selecting accelerograms for assessing the nonlinear demand of structures is based on first disaggregating (McGuire 1995), by causative magnitude and distance (M and R), the site's probabilistic seismic hazard analysis (PSHA) for the level of spectral acceleration (at a period near that of the first mode of the structure) at a specified probability (say, a 2% chance of exceedance in 50 years). The records are then chosen to match within tolerable limits the mean or modal value of the M and R, i.e., the expected value or most likely value of these characteristics given that exceedance. The records may also be selected for the expected style of faulting type and soil type, but we shall focus on M and R. Finally, the records are usually scaled to match in some average way the uniform hazard spectrum (UHS) or, as it is often recommended, precisely to the UHS level at a period near that of the first period of the structure when the structure is known (Shome et al. 1998). Several observations can be made about this procedure. For example, it is an unstated but implicit assumption that all this care is taken about the selected records' earthquake properties (e.g., M and R) because they (may) matter to linear or nonlinear response. But little information on this effect is available from earthquake engineers to pass on to the seismologist responsible for the selection. Lack of knowledge of the influence of seismological parameters on the structural response has driven the seismologists to be prudent and assume that all features (magnitude, faulting style, etc.) matter to structural response, and so they do their best to provide records accordingly.

The question of how best to select records is equivalent to asking, What earthquake parameters do we have to try to match when selecting the records? The concept of parsimony in engineering practice implies that the easiest way to try to answer this question is by first assuming that "it doesn't matter," which is equivalent to saying that the choice of records is a non-issue. Then, whether and under what conditions this assumption cannot be sustained is evaluated by a large number of examples and cases studies. In this framework, several structural types are considered belonging to both single-degree-of-freedom (SDOF) and multi-degree-of-freedom (MDOF) systems. SDOF systems are chosen to vary across a range of periods, backbones, and target ductilities. The MDOF systems belong to moderate period structures and have been chosen to represent quite different structural configurations. They include older reinforced concrete structures and steel moment-resisting frames with brittle connections.

The following procedure is used to test the importance of considering M and R when selecting records. First, a "target" group of sets is selected from a narrow magnitude-distance (M-R scenario) bin of available records. Then three size 10 samples of target sets are selected; each has two subsets representing the two horizontal components, yielding a total of six samples of size 10. The target sets scenario is a comparatively high-M, small-R one. To minimize potential directivity effects all values of R are greater than 15 km. For each structure considered, these records in the sample target sets are scaled to their overall median spectral acceleration at the first-mode period. This intra-bin scaling has been shown to be a good practice with respect to reducing the variance of the results of nonlinear analyses without introducing bias (Shome et al. 1998). The reduced variance increases the power of the statistical test to follow.

Second, another group of records, referred to as "arbitrary sets," is considered. As will be discussed in more detail below, these arbitrary sets are characterized by having been chosen (almost) at random with respect to the same features, M and R, which were carefully considered in the target sets. Five size 10 samples are selected, the two horizontal components yielding a total of 10 arbitrary samples of size 10. The arbitrary sets are also scaled to the common median first-mode period spectral acceleration of the target sets in order to mimic how any selected set of records might be scaled to the design target response spectrum.

Third, the structure in question is subjected to a nonlinear dynamic analysis under each of the many records. Median responses are estimated for all 6 plus 10 (16) record sets. The median of each of the arbitrary sets is compared with each of the target sets (6 times 10, or 60 comparisons). The comparison of medians is statistical and performed by a simple, conventional hypothesis test (Benjamin and Cornell 1970). Consistent with the assumption that "record selection doesn't matter," the null hypothesis is that the ratio of arbitrary-set median response to target-set median response is unity, i.e., that the medians are equal.

The question of the effect of scaling proposed record sets (such as one of these arbitrary sets) to the desired level (e.g., that of the median of the target sets here) was addressed next. The 10 arbitrary sets did not require a degree of scaling significantly greater than 1 to reach the median of the six targets above. Therefore, another group of stronger target sets was constructed for this phase of the study. These were selected from records obtained within 15 km. As described below; care was taken to avoid records with significant directivity effects. In all other respects the same three steps above were repeated. The null hypothesis is, again, that such scaling “does not matter” (i.e., that median response to a scaled arbitrary set is the same as that median response to the target set).

CLASSES AND RECORDS

All the records used came from the Pacific Earthquake Engineering Research Center (PEER) database (http://peer.berkeley.edu/smcat/), ensuring uniform processing. However, all the accelerograms in both of the groups of sets have been selected with some boundary conditions in order to better reduce the influence of those factors that are not in the objective of the study. In particular, only California events have been considered, recorded on NEHRP C-D soil class and coming from free-field or one-story building instrument housing. These features make the records definable as “ordinary,” avoiding site and housing response effects. Moreover, for addressing the selection issue, the records belong to the *far field* (defined here as closest distance to rupture greater than 15 km) in order to better avoid directivity pulse-type effects. Other features such as hanging/foot wall and fault mechanism are permitted to vary among the record sets considered, as they do not cause systematic peaks in the spectra. Next we address how the various record sets were selected from the “reduced” catalog defined above.

CLASS OF TARGET SETS

The target sets for the record selection study are designed to be representative of a specific scenario event (M and R) that might be the realistic threat at a particular site, here a moment magnitude 7 at 20 km, defined as closest distance to fault rupture. This target event was chosen to be as large and close as feasible, given the wish to have several samples of the target sets and given the limited number of large magnitude, close records in the catalog. (The records must also respect the general selection criteria presented just above.) In order to best represent what might occur in the future and to reduce correlation or “overlapping” due to event commonality, it is desirable to have the 10 records in each set come from 10 different events. This requirement conflicts with the desire to have a large target magnitude and to sample events close to the target in magnitude. The compromise was to use 5 events and 2 records per event. This decision led to 5 events with magnitude range 6.7 to 7.4. Starting from this point, 6 different sets of 10 records each have been arranged such that almost all the records are in the narrow distance range 20 ± 5 km. The comparatively small sample size of 10 events in each set has been chosen because ten is the order of magnitude of size used in recommended earthquake engineering practice (which is typically as small as three to seven) (ICC 2000) (The total, or pooled, set of all records will also be considered, but the breakdown into sets of a more conventional size is considered more representative and hence more instructive and transparent.) Further, no two target sets have more than 1 record in common out of the 10; complete avoidance of overlap was not feasible because not all five events had the 12 records necessary to fill out the six target sets within or near the distance range. These selection limitations on events and records are designed to make the sets as nearly independent as possible given the limited number of records available.

The records in the target sets are named T1a, T1b, T2a, T2b, T3a, and T3b. Sets with a common number such as T1a and T1b contain components from precisely the same 10 three-component recordings. The "a" and "b" refer to the fact that one horizontal component is used in "a" and the other in "b." In Table A1 (see Appendix), records in T1a, T2a, and T3a are listed so that all target sets can easily be retrieved in the PEER online database.

CLASS OF ARBITRARY SETS

These sets were chosen effectively randomly from the catalog without regard to magnitude or distance subject only to the general constraints presented above (California, soil type, and distance). The *arbitrary sets* are 10 sets of 10 records each.[1] They come from California events in a comparatively wide range of moderate magnitudes ($6.4 < M < 7.4$) and distances ($15 < R < 50$ km). The records in each set are chosen randomly (*without* replacement) first from the list of events, ensuring 10 different events in this case, and then from the available distances within each selected event to the degree possible; because of limits on the number of recordings/distances in such event, it was necessary to have two records from one event in some cases in order to construct 10 arbitrary sets. The upper bound is the limit of the catalog; the lower bound of 6.4 was selected because it is a full magnitude unit below the upper value and because there are sufficient events and records in the catalog that in practical application one need go no lower than this to have a large sample from which to select the relatively few records (3 to 10) that are needed in an application.

The sets in the arbitrarily chosen group are named from A1a, A1b to A5a, A5b. The "a" and "b" apply as above. Records belonging to A1a–A5a are reported in Table A2 (see Appendix).

In short, unusual care has been taken in selecting records in the A and T sets to make the individual samples within the A and T sets as nearly random and exclusive (non-overlapping) as possible. The A and T sets, of course, overlap one another. Details about these sets' M-R differences and similarities may be found elsewhere (Iervolino 2004); they are also discussed in the conclusions section below.

DESIGN OF ANALYSES AND CASE STUDIES

To establish the validity of the (over) simply stated hypothesis that "it does not matter how one selects records," a series of structures have to be chosen. In order to make the conclusions of the study broad, wide-ranging cases have been considered. The different structural features considered to be most meaningful to be investigated are as follows: (1) first natural period, (2) force-deformation or hysteresis relationship, (3) target ductility, (4) number of degrees of freedom, and (5) structural type (concrete or steel). For each of these factors a wide range has been considered in order to help establish the limits of acceptance of the hypothesis.

[1] One of the A sets (A5) is actually made of nine accelerograms since one of the chosen stations has only one recorded component; A5b is therefore made up of only nine elements.

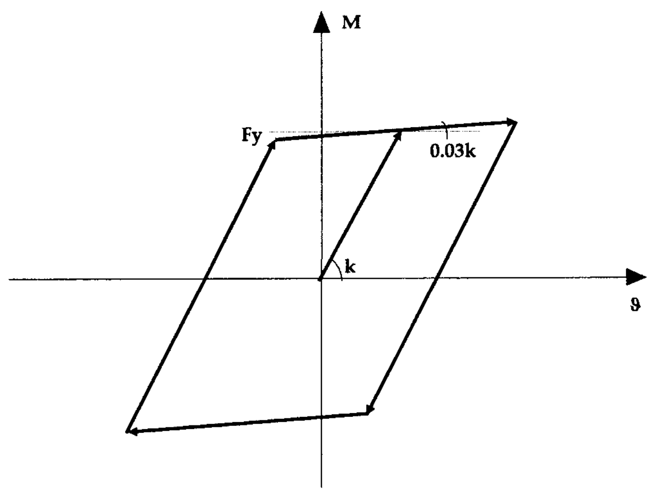

Figure 1. SDOF bilinear hysteresis.

SINGLE-DEGREE-OF-FREEDOM SYSTEMS

Three different periods SDOF systems have been considered: very short (0.1 sec), moderate (1.5 sec), and very long (4 sec), in order to investigate if conclusions reached at moderate periods seem to hold at extreme periods. Most nonlinear SDOF system study is based on simple bilinear systems with a second stiffness equal to 3% of the first; see Figure 1 for a hysteresis rule example.

For each of the three periods, two yield strengths are selected to give median duc-

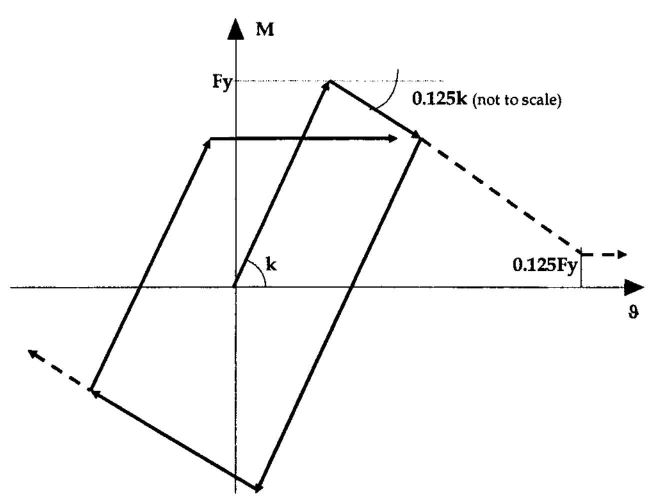

Figure 2. SDOF strength-degrading trilinear hysteresis.

tility of approximately 2 and 6 under the target records. Viscous damping is always 5% of critical. For the most interesting moderate period case (T=1.5 sec case), a second trilinear backbone with a degrading strength branch (see Figure 2) is considered, again with two strength levels. Table 1 summarizes the SDOF structural configurations analyzed under both the target and the arbitrary groups of sets.

MULTIPLE-DEGREE-OF-FREEDOM SYSTEMS

Both of the MDOF structural models are moment-resisting frames, one of reinforced concrete (RC) and one of steel. The former is an older, nonductile RC building in Van Nuys, California, studied as part of the PEER testbed program (Porter et al. 2002), while the other structure is the nine-story Los Angeles SAC steel building (Gupta and Krawinkler 1999) with brittle connections (Luco and Cornell 2000).

The reinforced concrete structure is modeled with strength-degrading moment and shear behavior in the nonlinear range of the member-end hinges. The Van Nuys building is a seven-story, 6,200-m^2 (66,000-sf) hotel located in the San Fernando Valley (Southern California). The hotel was built in 1966; its design refers to the 1964 Los Angeles City Building Code. The plan view of the building is rectangular, 21 m (63 ft) by 50 m (150 ft), three bays by eight bays, seven stories tall (see Figure 3). The reinforcing steel lacks modern ductile detailing. Moment frames along the perimeter provide the primary seismic force resistance. See the Van Nuys Testbed Committee (2002) for details. A two-dimensional (2-D) model of the transverse frame has analyzed here by DRAIN-2D software (Jalayer 2003). Its transverse vibration has a dominant period of about 0.85 sec.

The other analyzed MDOF is a 2-D model of the SAC LA-9 frame. It is one of the structures considered as part of the SAC steel project, designed to the 1994 *Uniform Building Code* (Figure 4). The nine-story benchmark structure is 45.73 m (150 ft) by 45.73 m (150 ft) in plan, and 37.19 m (122 ft) in elevation. The bays are 9.15 m (30 ft) on center, in both directions, with five bays each in the north-south and east-west directions (Gupta and Krawinkler 1999). The first natural period of this building is between 1.5 and 2 sec, corresponding to the moderate period SDOF already considered. The fracture of the "pre-Northridge" brittle connections is introduced into the DRAIN-2DX program using rotational spring elements at the ends of each elastic beam. The element was developed by Foutch and Shi (2002).

ANALYSES

The SDOF and MDOF cases presented have been analyzed with the groups of sets described as input. The considered response parameter is the peak in-time drift for the

Table 1. SDOF cases

Period	T=1.5 sec				T=0.1 sec		T=4 sec	
Backbone	Bilinear		Trilinear		Bilinear		Bilinear	
Ductility	2	6	2	6	2	6	2	6

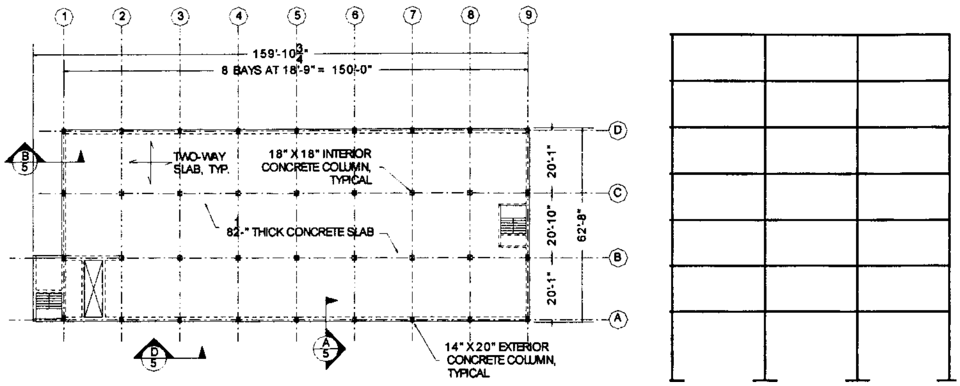

Figure 3. Van Nuys Holiday Inn plan and transverse elevation (http://www.peertestbeds.net).

SDOF cases and the maximum interstory peak drift over all the stories for the MDOF cases. The latter is used in recent criteria for frame structures such as the SAC steel project (see, e.g., Gupta and Krawinkler 1999) as an indicator of the extreme rotation demands in the joints and of possible collapse due to global instability. We shall refer to these as simply "drift" in what follows.

Each of the considered structural cases has been analyzed first by running the 6 different considered target sets as recorded, meaning without any scaling. For each structural model this produces a "cloud" of points (see Figure 5) in the plane of drift, θ, (the selected Engineering Demand Parameter or EDP) versus first-mode spectral acceleration, Sa. The records belonging to the target sets have then been scaled to the median first-mode spectral acceleration of all the 6 target sets (estimated by the geometric mean)

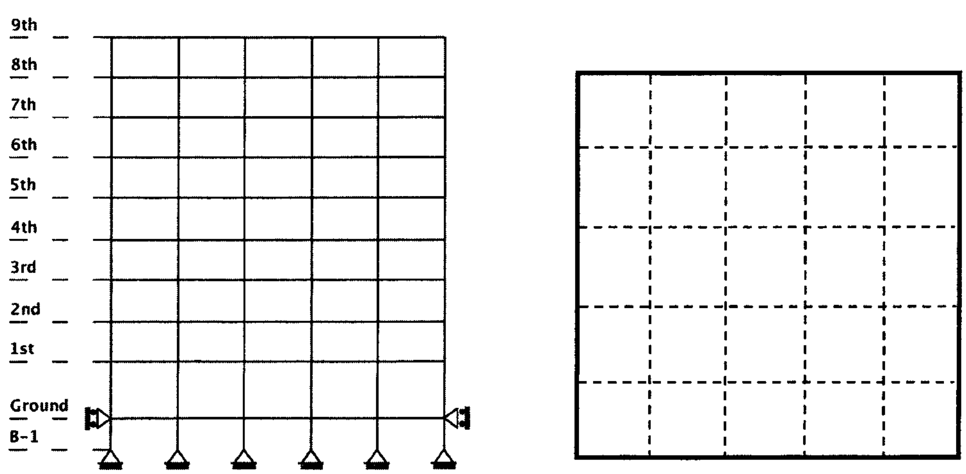

Figure 4. SAC LA-9 story elevation and plan (Gupta and Krawinkler 1999).

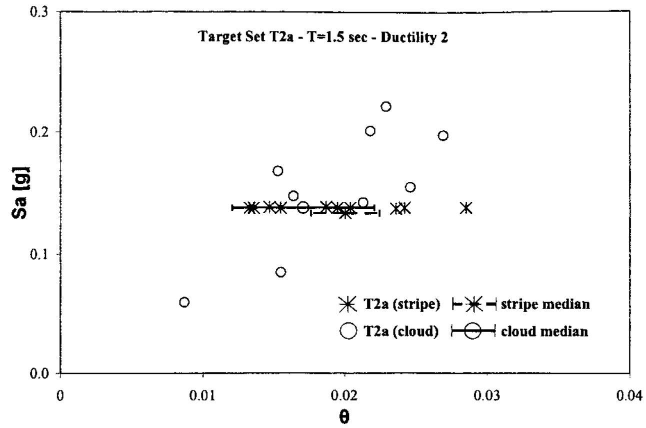

Figure 5. Cloud and striping of a target set to the median Sa for T=1.5 sec SDOF–ductility 2 case. Note the similarity of the two drift medians and the smaller dispersion (proportional to horizontal bar dimension) in the stripe.

(Figure 5). This second "stripe" scaling is not necessary, but it gives virtually the same median EDP (as shown by Shome et al. [1998] and as confirmed here by these two steps). It also gives smaller dispersion and hence provides for stronger significance tests. The arbitrary sets have been analyzed in the same way. Finally, to permit the comparison of the target and the arbitrary sets, the records within the arbitrary sets have all been scaled to the median spectral acceleration of the targets sets (see Figure 6). The medians of the response parameters (maximum drift) of the two different sets are next compared statistically for equality (or not).

The hypothesis that it does not matter how one chooses records means that the results of the arbitrary sets scaled to the median spectral accelerations of the target should give virtually the same results in terms of the chosen response parameter. This equivalence for each structural case can be assessed by statistically testing the ratio of the estimated median of the drift results of each of the arbitrary sets Aij ($i=1,2,\ldots,5$; $j=a,b$) divided by the estimated median of the results of each of the target sets Tij ($i=1,2,3$; $j=a,b$).

In the following relation the ratio of the estimated medians of the drift responses of arbitrary set, x, and target set, y, is defined as z:

$$z=\frac{\bar{\theta}_x}{\bar{\theta}_y} \tag{1}$$

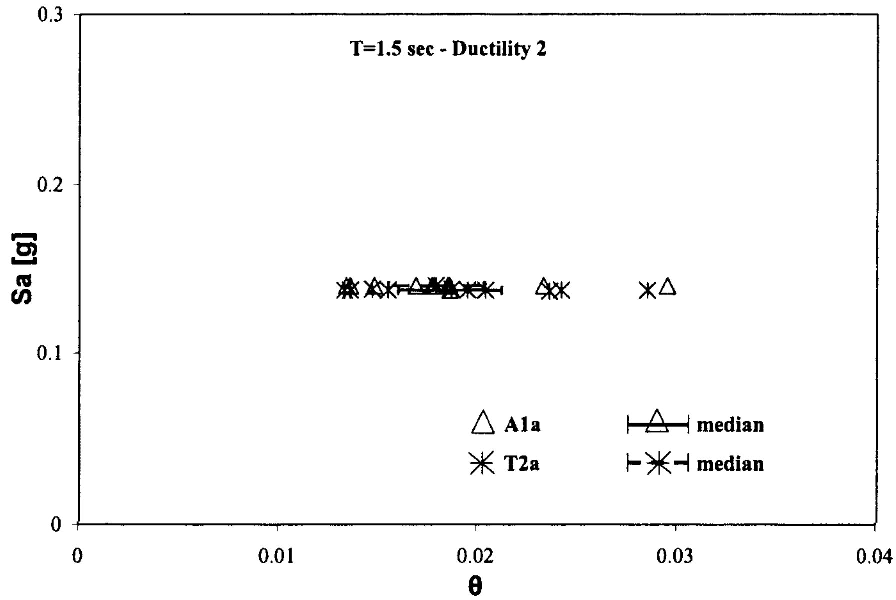

Figure 6. An arbitrary set and a target set striped to the same Sa (median of target sets) for T = 1.5 sec SDOF–ductility 2 case. Note the similarity of the two drift medians and comparable dispersion.

in which $\bar{\theta}_x$ and $\bar{\theta}_y$ are, respectively, the estimated medians of set x and set y. Recall that the estimate, which is the geometric mean, is also equal to log base e raised to a power equal to the average of the natural logs of the drift.

The estimation of standard error of Z can be evaluated as:

$$\beta_z = \sqrt{\frac{\beta^2{}_x}{n_1} + \frac{\beta^2{}_y}{n_2}} \qquad (2)$$

where n_1 and n_2 are the sample sizes of the particular arbitrary set and target set, respectively; β_x and β_y are the standard deviations of the natural logarithms of the particular arbitrary set and target set, respectively.

Under the assumption that the individual responses are lognormal (which has been reported by Shome et al. (1998) and others) the natural log of the ratio in Equation 1 divided by β_Z is distributed as a student-t with (n_1+n_2-2) degrees of freedom. This number is typically 18 here. Statistically speaking, the assumption being tested is that the medians of the probability distribution of the drift results are the same for both the target and the arbitrary sets. Or equivalently, that the means of the natural logs are equal. The null hypothesis, H_0, of the test is

$$H_0\text{:}\mathit{median\ responses\ are\ equal}$$

To the acceptance of this hypothesis we can associate a statistical significance level, which corresponds to the risk of rejecting when it is, in fact, correct.

$$\mathit{Significance\ level} = P[\mathit{reject}\ H_0 | H_0\ \mathit{correct}] \tag{3}$$

Two significance levels have been considered in the present study; they correspond to the "1.5-sigma and 2-sigma" levels (here the "1.5-β and 2-β" levels). For the two-sided student-t PDF with 18 degrees of freedom (DOF) these levels correspond to 0.152 and 0.0608, respectively (these values differ only slightly from those for the simple Gaussian distribution, which the t-distribution approaches for large DOF values). Therefore, to test whether one can accept the hypothesis that a pair of target and arbitrary set medians are equal, simply calculate the natural log of their ratio (ln(z)), divide that by the estimated standard error, β_Z, (Equation 2) and confirm whether that result is "close" to zero, where close means within ±1.5 or ±2.0 for the 15% and 6% significance levels, respectively. For example, the upper left-hand case in Table 2 yields a ratio of 1.07 whose natural log is about 0.07; the standard error is about 0.14. The former divided by the latter is 0.48 (to two significant figures). This value is less than 1.5 (or 2.0), so the equality of medians hypothesis may be accepted at the 15% (and 6%) significance level. Note that if the ratio is close to unity, an approximate check is to compare simply the deviation from unity (1.07−1.0 or 0.07 in this example) with 1.5 (or 2.0) times β (0.21 or 0.28 in this example). This is adequate in most cases here. Note that even if the hypothesis is true one would expect to reject it (incorrectly) in about 15% (or 6%) of the cases.

Finally before making the arbitrary to target set comparisons, the target sets have been compared to confirm that they are "equivalent." This means that the target set responses have been statistically compared among themselves that the hypothesis of equality of medians has been accepted.

RESULTS AND DISCUSSION

In the following tables, Tables 2–6, all the case study results are summarized. Each table shows a particular structural case; in each cell are the median ratios, z, for the correspondent sets and, below, the standard error estimation, β_Z, in italics. Those cases for which the null hypothesis is rejected at the $2-\beta$ or 6% significance level are shown in bold font. (Those rejected at the 15% level are shown underlined.) For example, for the bilinear T = 1.5 sec, ductility 6 (Table 2) case with arbitrary set A5a and target set T3b, the ratio of the median of the arbitrary set responses to that of the target set is shown to be 0.69. It is statistically different from one because the log of 0.69 divided by 0.15 is −2.47, which is less than −2. Hence it is shown in bold. Note the ratio to its left is smaller (0.66) but the standard error estimation is larger (0.18), so in this test the medians are not different at the 6% level (although they are at 15% level).

Table 2. Ratio of medians drifts and standard errors for the T=1.5 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	T1a	T2a	T3a	T1b	T2b	T3b	T	$\mu \approx 2$	T1a	T2a	T3a	T1b	T2b	T3b	T
A1a	1.07	0.88	1.10	1.00	0.86	0.89	0.96	A1a	0.96	0.89	0.93	0.95	0.93	0.93	0.93
	0.14	*0.17*	*0.15*	*0.17*	*0.16*	*0.13*	*0.10*		*0.11*	*0.10*	*0.11*	*0.09*	*0.09*	*0.10*	*0.08*
A2a	1.27	1.05	1.31	1.18	1.02	1.06	1.14	A2a	1.09	1.02	1.06	1.08	1.05	1.06	1.06
	0.18	*0.21*	*0.19*	*0.21*	*0.20*	*0.18*	*0.16*		*0.12*	*0.11*	*0.11*	*0.10*	*0.09*	*0.11*	*0.09*
A3a	1.12	0.93	1.15	1.05	0.90	0.93	1.01	A3a	1.02	0.95	0.99	1.01	0.98	0.99	0.99
	0.16	*0.18*	*0.17*	*0.18*	*0.18*	*0.15*	*0.12*		*0.10*	*0.08*	*0.09*	*0.07*	*0.06*	*0.08*	*0.05*
A4a	1.06	0.87	1.09	0.99	0.85	0.88	0.95	A4a	0.99	0.92	0.96	0.98	0.96	0.96	0.96
	0.15	*0.17*	*0.16*	*0.18*	*0.17*	*0.14*	*0.11*		*0.10*	*0.09*	*0.10*	*0.08*	*0.07*	*0.09*	*0.07*
A5a	1.09	0.90	1.12	1.02	0.88	0.91	0.98	A5a	0.99	0.92	0.96	0.98	0.95	0.96	0.96
	0.16	*0.18*	*0.17*	*0.18*	*0.18*	*0.15*	*0.12*		*0.11*	*0.10*	*0.10*	*0.09*	*0.08*	*0.10*	*0.07*
A1b	1.22	1.01	1.25	1.14	0.98	1.01	1.10	A1b	1.06	0.98	1.02	1.05	1.02	1.03	1.03
	0.15	*0.18*	*0.17*	*0.18*	*0.18*	*0.15*	*0.12*		*0.10*	*0.09*	*0.10*	*0.08*	*0.07*	*0.09*	*0.06*
A2b	1.16	0.96	1.20	1.09	0.93	0.97	1.05	A2b	1.08	1.00	1.04	1.07	1.04	1.05	1.04
	0.23	*0.25*	*0.24*	*0.25*	*0.25*	*0.23*	*0.21*		*0.13*	*0.12*	*0.12*	*0.11*	*0.11*	*0.12*	*0.10*
A3b	1.15	0.95	1.18	1.07	0.92	0.96	1.03	A3b	1.04	0.96	1.00	1.03	1.00	1.01	1.01
	0.17	*0.19*	*0.18*	*0.19*	*0.19*	*0.17*	*0.14*		*0.10*	*0.09*	*0.10*	*0.08*	*0.07*	*0.09*	*0.07*
A4b	1.18	0.98	1.22	1.11	0.95	0.99	1.07	A4b	1.06	0.98	1.02	1.05	1.02	1.02	1.02
	0.16	*0.18*	*0.17*	*0.18*	*0.18*	*0.15*	*0.13*		*0.11*	*0.09*	*0.10*	*0.08*	*0.08*	*0.09*	*0.07*
A5b	0.83	**0.68**	0.85	0.77	**0.66**	**0.69**	**0.75**	A5b	0.95	0.88	0.92	0.94	0.92	0.92	0.92
	0.15	*0.18*	*0.16*	*0.18*	*0.18*	*0.15*	*0.12*		*0.10*	*0.09*	*0.10*	*0.08*	*0.08*	*0.09*	*0.07*
A	1.11	0.92	1.14	1.04	0.89	0.93	1.00	A	1.02	0.95	0.99	1.01	0.99	0.99	0.99
	0.11	*0.15*	*0.13*	*0.15*	*0.14*	*0.11*	*0.07*		*0.09*	*0.07*	*0.08*	*0.06*	*0.05*	*0.07*	*0.02*

In the last column of each table, the 6 target sets are pooled to create a single large target set ($n_1=52$ is the number of *distinct* records in the target sets) to compare to each of the arbitrary sets, while in the last row of each table the 10 arbitrary sets are merged ($n_2=99$) and compared to each target set. Finally in the lower right-hand corner of each table both the target set and the arbitrary sets are merged and compared. (Note standard errors are smaller on the edges.)

Consider Table 2 for the moderate period bilinear SDOF system. Looking first at the pooled (large sample) results in the lower right-hand corner of each half of the table, it is clear that the arbitrarily chosen records are indeed virtually indistinguishable from those of the target scenario with respect to their impact on the nonlinear response of this oscillator. The observed ratios are virtually unity (0.99 and 1.00) with standard errors of 2 and 7% for ductility 2 and 6, respectively. The full table of 120 ratios, however, gives a better sense of what degree of confirmation one can expect from a professionally practical sample size. The observed ratios scatter either side of unity, ranging from as low as

Table 3. Ratio of medians drifts and standard errors for the T=1.5 sec trilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	T1a	T2a	T3a	T1b	T2b	T3b	T	$\mu \approx 2$	T1a	T2a	T3a	T1b	T2b	T3b	T
A1a	0.72	0.57	0.97	0.57	0.80	0.61	0.69	A1a	1.02	0.96	1.06	0.93	1.06	1.03	1.01
	0.40	*0.41*	*0.42*	*0.43*	*0.35*	*0.44*	*0.30*		*0.21*	*0.12*	*0.12*	*0.16*	*0.17*	*0.16*	*0.11*
A2a	1.93	1.54	2.60	1.53	2.16	1.65	1.87	A2a	1.26	1.19	1.30	1.15	1.31	1.28	1.25
	0.52	*0.52*	*0.53*	*0.54*	*0.47*	*0.55*	*0.44*		*0.26*	*0.19*	*0.19*	*0.22*	*0.23*	*0.22*	*0.19*
A3a	1.21	0.97	1.63	0.96	1.35	1.03	1.17	A3a	1.10	1.04	1.14	1.00	1.15	1.12	1.09
	0.39	*0.40*	*0.41*	*0.42*	*0.33*	*0.43*	*0.28*		*0.23*	*0.15*	*0.15*	*0.18*	*0.19*	*0.19*	*0.14*
A4a	1.01	0.80	1.36	0.80	1.13	0.86	0.97	A4a	1.02	0.96	1.05	0.93	1.06	1.03	1.01
	0.40	*0.40*	*0.41*	*0.42*	*0.34*	*0.44*	*0.29*		*0.23*	*0.15*	*0.15*	*0.18*	*0.19*	*0.19*	*0.14*
A5a	1.32	1.05	1.78	1.05	1.47	1.13	1.27	A5a	1.09	1.02	1.12	0.99	1.13	1.10	1.08
	0.48	*0.47*	*0.48*	*0.50*	*0.42*	*0.50*	*0.38*		*0.21*	*0.12*	*0.12*	*0.16*	*0.17*	*0.17*	*0.12*
A1b	0.86	0.69	1.16	0.68	0.96	0.73	0.83	A1b	1.05	0.99	1.09	0.96	1.09	1.06	1.04
	0.39	*0.39*	*0.40*	*0.41*	*0.33*	*0.43*	*0.28*		*0.20*	*0.10*	*0.10*	*0.15*	*0.16*	*0.15*	*0.09*
A2b	1.35	1.08	1.83	1.08	1.52	1.16	1.31	A2b	1.35	1.28	1.40	1.23	1.41	1.37	1.34
	0.50	*0.50*	*0.51*	*0.52*	*0.45*	*0.53*	*0.42*		*0.32*	*0.27*	*0.27*	*0.29*	*0.30*	*0.29*	*0.27*
A3b	1.99	1.59	**2.68**	1.58	**2.22**	1.70	**1.92**	A3b	0.95	0.90	0.98	0.87	0.99	0.97	0.94
	0.42	*0.43*	*0.43*	*0.44*	*0.37*	*0.46*	*0.32*		*0.21*	*0.12*	*0.12*	*0.16*	*0.17*	*0.16*	*0.11*
A4b	1.51	1.21	2.04	1.20	1.69	1.29	1.46	A4b	1.00	0.94	1.03	0.91	1.04	1.01	0.99
	0.45	*0.45*	*0.46*	*0.47*	*0.39*	*0.48*	*0.35*		*0.22*	*0.14*	*0.14*	*0.18*	*0.19*	*0.18*	*0.14*
A5b	1.02	0.81	1.37	0.81	1.14	0.87	0.98	A5b	0.92	0.86	0.95	0.83	0.95	0.93	0.91
	0.42	*0.42*	*0.43*	*0.44*	*0.36*	*0.46*	*0.32*		*0.23*	*0.16*	*0.16*	*0.19*	*0.20*	*0.19*	*0.15*
A	1.23	0.98	*1.66*	0.98	1.38	0.76	1.19	A	1.07	1.01	1.11	0.98	1.11	0.87	1.06
	0.31	*0.32*	*0.33*	*0.34*	*0.24*	*0.43*	*0.16*		*0.19*	*0.08*	*0.07*	*0.13*	*0.15*	*0.28*	*0.07*

0.66 to as high as 1.31 for the ductility 6 case. Typically they are within 20% of one for that case and within 10% for the ductility 2 case. The standard errors of estimation of the ratio are typically estimated to be about 0.1 for the ductility 2 case, and 0.2 for the ductility 6 case. Note that both the numerators and the denominators of the ratios suffer from statistical variability. (It is worth noting that Equation 2 implies that for this problem the individual (log) standard deviations, β_x and β_y, for an individual record are about $\sqrt{5}$ or 2.2 times the standard error of estimation of the small set ratios. Hence, in this moderate-period case, the individual record-to-record log standard deviations of the displacements are about +/−40% for ductility 6 and +/−20% for ductility 2.)

None of the 60 ductility 2 cases and three of the 60 cases of ductility 6 have median ratios found to be statistically significantly different from one (at the 6% significance level implied by the ± 2-β bounds). This total of four is about 7% of the total 60 cases. The other cases are similar. For example, the T=1.5 trilinear case (with strength degradation) (Table 3) and the 0.1 sec bilinear case (Table 4) both have only 2 of 60 ductility 6 sets that are statistically significant. The T=4 sec case (Table 5) has 7 of 30 at ductility 6, and 4 of 30 at ductility 2. These fractions are somewhat different from the 6% fraction that might be expected if the hypothesis of unit ratio is in fact true, i.e., if the response

Table 4. Ratio of medians drifts and standard errors for the T=0.1 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu\approx 6$	T1a	T2a	T3a	T1b	T2b	T3b	T	$\mu\approx 2$	T1a	T2a	T3a	T1b	T2b	T3b	T
A1a	2.04	1.50	1.44	1.42	1.26	1.37	1.49	A1a	1.23	1.18	1.07	1.20	1.08	1.13	1.15
	0.38	*0.37*	*0.35*	*0.37*	*0.37*	*0.34*	*0.29*		*0.17*	*0.16*	*0.18*	*0.19*	*0.22*	*0.18*	*0.15*
A2a	1.38	1.02	0.97	0.96	0.85	0.93	1.01	A2a	1.14	1.09	0.98	1.11	1.00	1.04	1.06
	0.41	*0.39*	*0.37*	*0.40*	*0.39*	*0.37*	*0.32*		*0.13*	*0.13*	*0.15*	*0.16*	*0.20*	*0.15*	*0.11*
A3a	**2.21**	1.62	1.55	1.54	1.36	1.48	1.61	A3a	1.62	1.55	1.40	1.58	1.42	1.48	*1.51*
	0.38	*0.37*	*0.35*	*0.38*	*0.37*	*0.35*	*0.30*		*0.25*	*0.25*	*0.26*	*0.26*	*0.29*	*0.26*	*0.24*
A4a	1.70	1.25	1.20	1.18	1.04	1.14	1.24	A4a	1.45	1.39	1.25	1.41	1.27	1.33	1.35
	0.43	*0.41*	*0.40*	*0.42*	*0.41*	*0.39*	*0.35*		*0.22*	*0.21*	*0.23*	*0.23*	*0.26*	*0.23*	*0.20*
A5a	0.95	0.70	0.67	0.66	0.59	0.64	0.69	A5a	1.03	0.99	0.89	1.01	0.90	0.95	0.96
	0.39	*0.38*	*0.36*	*0.38*	*0.38*	*0.35*	*0.29*		*0.12*	*0.11*	*0.14*	*0.15*	*0.19*	*0.14*	*0.09*
A1b	2.04	1.50	1.44	1.42	1.25	1.37	1.48	A1b	1.51	1.44	1.30	1.47	1.32	1.38	1.40
	0.42	*0.41*	*0.39*	*0.42*	*0.41*	*0.39*	*0.35*		*0.26*	*0.25*	*0.27*	*0.27*	*0.30*	*0.27*	*0.25*
A2b	1.34	0.99	0.95	0.93	0.83	0.90	0.98	A2b	0.99	0.95	0.86	0.97	0.87	0.91	0.92
	0.39	*0.38*	*0.36*	*0.38*	*0.37*	*0.35*	*0.31*		*0.14*	*0.13*	*0.16*	*0.16*	*0.20*	*0.16*	*0.12*
A3b	2.21	1.62	1.56	1.53	1.35	1.48	1.60	A3b	1.39	1.34	1.20	1.36	1.22	1.28	1.30
	0.40	*0.39*	*0.37*	*0.39*	*0.39*	*0.37*	*0.32*		*0.22*	*0.22*	*0.23*	*0.24*	*0.26*	*0.23*	*0.21*
A4b	**2.17**	1.59	1.53	1.51	1.33	1.45	1.58	A4b	1.19	1.15	1.03	1.17	1.05	1.10	1.11
	0.36	*0.34*	*0.32*	*0.35*	*0.34*	*0.31*	*0.26*		*0.21*	*0.20*	*0.22*	*0.22*	*0.25*	*0.22*	*0.19*
A5b	1.02	0.75	0.72	0.71	0.63	0.69	0.74	A5b	0.98	0.94	0.84	0.95	0.85	0.90	0.91
	0.34	*0.31*	*0.29*	*0.33*	*0.31*	*0.29*	*0.23*		*0.11*	*0.10*	*0.13*	*0.14*	*0.18*	*0.13*	*0.08*
A	1.64	1.20	1.16	1.14	1.01	1.05	1.19	A	**1.24**	1.19	1.07	1.21	1.08	1.14	1.15
	0.28	*0.25*	*0.22*	*0.26*	*0.25*	*0.23*	*0.14*		*0.10*	*0.09*	*0.12*	*0.13*	*0.18*	*0.13*	*0.07*

results for the arbitrarily selected records are the same as those from the {Magnitude =7, R=20 km} target sets. Only the ductility 2 T=0.1 sec case is statistically significant when pooled to the largest sample size. (Its test value is 2.00 based on a median ratio of 1.15 and β of 0.07.) Considering all former SDOF cases, six of the 16×4 or 64, or 9%, of the pooled columns and rows suggest rejection. Study of the patterns within the tables shows that individual rows or columns (sets) tend to dominate. For example, all seven of the ductility 6 T=4 sec rejected cases are associated with just two target sets, T2a and T3a, suggesting that these two sets may be (randomly) "stronger" than typical. Yet three arbitrary sets, A5a, A5b, and A2b dominate the ductility 2 rejections, the latter two being "stronger." These effects are reflected in the column and row pools. The large pooled set is not significantly different for one, however. Similar, weaker such effects are seen in the other cases. Notice, however, the "strong" or "weak" sets may not carry over from one ductility to another. In sum, these patterns imply that the 60 small sample ratios are not fully independent of one another as they share common sets. Hence one unusual set in 10 (or 6) may cause multiple "failures" within the 60 cases.

The MDOF cases are both quite benign, that is, there is only one statistically sig-

Table 5. Ratio of medians drifts and standard errors for the T=4 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu\approx 6$	T1a	T2a	T3a	T1b	T2b	T3b	T	$\mu\approx 2$	T1a	T2a	T3a	T1b	T2b	T3b	T
A1a	0.89	**0.66**	**0.71**	0.98	0.79	0.82	**0.80**	A1a	0.94	0.83	0.84	1.01	0.93	0.99	0.92
	0.14	*0.15*	*0.13*	*0.13*	*0.17*	*0.14*	*0.09*		*0.09*	*0.10*	*0.09*	*0.08*	*0.12*	*0.07*	*0.06*
A2a	1.14	0.85	0.91	1.26	1.02	1.06	1.03	A2a	1.12	0.99	1.01	1.21	1.11	1.18	1.10
	0.17	*0.18*	*0.17*	*0.16*	*0.20*	*0.17*	*0.13*		*0.11*	*0.11*	*0.10*	*0.10*	*0.13*	*0.09*	*0.09*
A3a	1.00	0.74	0.79	1.10	0.89	0.92	0.90	A3a	0.96	0.84	0.86	1.03	0.94	1.01	0.94
	0.15	*0.17*	*0.15*	*0.14*	*0.18*	*0.16*	*0.11*		*0.10*	*0.10*	*0.10*	*0.09*	*0.12*	*0.08*	*0.07*
A4a	1.04	0.77	0.83	1.15	0.93	0.96	0.94	A4a	1.03	0.91	0.93	1.11	1.02	1.08	1.01
	0.16	*0.18*	*0.16*	*0.15*	*0.19*	*0.17*	*0.12*		*0.09*	*0.10*	*0.09*	*0.08*	*0.12*	*0.07*	*0.06*
A5a	0.96	0.71	0.76	1.06	0.85	0.88	0.86	A5a	0.90	**0.79**	**0.81**	0.97	0.89	0.94	*0.88*
	0.17	*0.18*	*0.16*	*0.16*	*0.19*	*0.17*	*0.12*		*0.11*	*0.11*	*0.10*	*0.09*	*0.13*	*0.09*	*0.07*
A1b	1.05	0.77	0.83	1.16	0.93	0.97	0.94	A1b	1.10	0.97	0.99	1.19	1.09	1.16	1.08
	0.15	*0.16*	*0.14*	*0.14*	*0.18*	*0.15*	*0.10*		*0.10*	*0.11*	*0.10*	*0.09*	*0.12*	*0.08*	*0.07*
A2b	1.19	0.88	0.95	1.32	1.06	1.10	1.07	A2b	1.16	1.02	1.04	**1.25**	1.14	**1.22**	**1.14**
	0.16	*0.18*	*0.16*	*0.16*	*0.19*	*0.17*	*0.12*		*0.10*	*0.10*	*0.09*	*0.08*	*0.12*	*0.07*	*0.06*
A3b	0.88	**0.65**	**0.70**	0.98	0.79	0.82	**0.80**	A3b	0.98	0.87	0.88	1.06	0.97	1.03	0.96
	0.14	*0.15*	*0.14*	*0.13*	*0.17*	*0.15*	*0.09*		*0.10*	*0.10*	*0.09*	*0.08*	*0.12*	*0.07*	*0.07*
A4b	0.95	**0.70**	0.75	1.05	0.84	0.88	0.85	A4b	0.93	0.82	0.84	1.00	0.92	0.98	0.91
	0.15	*0.17*	*0.15*	*0.14*	*0.18*	*0.16*	*0.11*		*0.12*	*0.13*	*0.12*	*0.11*	*0.14*	*0.11*	*0.10*
A5b	0.88	**0.65**	**0.70**	0.97	0.78	0.81	**0.79**	A5b	0.99	0.88	0.89	1.07	0.98	1.04	0.97
	0.16	*0.17*	*0.15*	*0.15*	*0.19*	*0.16*	*0.11*		*0.10*	*0.11*	*0.10*	*0.09*	*0.12*	*0.08*	*0.07*
A	0.99	**0.73**	0.79	1.10	0.88	0.92	0.90	A	1.01	0.89	0.91	1.09	1.00	1.06	0.99
	0.12	*0.14*	*0.12*	*0.11*	*0.16*	*0.13*	*0.06*		*0.08*	*0.08*	*0.07*	*0.06*	*0.11*	*0.05*	*0.04*

nificant case (at the 6% level) in the 60 small sample comparisons, two pooled row significant cases, and a total sample median ratio very close to one in both cases.

Note that only one of the total of 10 structural cases, the T=0.1 bilinear SDOF case (Table 4), shows a large (pooled) sample ratio statistically different from one. Its ratio of 1.15 is just larger than twice its standard error of estimation of 0.07; further, for this 0.1 sec case, the small sample cases show only one arbitrary set (A3a) with a number of significant ratios in the ductility 2 case, and just one target set (T1a) with several significant ratios in the ductility 6 case. Apparently, arbitrary set A3a is unusually severe and target set T1a is unusually weak.

It should be noted that, relative to the moderate period bilinear case, the trilinear and 0.1 sec cases have large standard errors (caused by large record-to-record variability), while that of the 4 sec case is similar to the 1.5 sec bilinear SDOF case in respect to the size of the standard errors. The MDOF systems have standard errors similar to that of the ductility 6 1.5-sec bilinear SDOF case. While, even for the largest pooled sample sizes, the observed mean ratios of the trilinear and 0.1-sec case are often quite different from unity (almost 20% in the ductility 6 cases), the important question is how different is the ratio relative to the standard error of estimation (less than 1.5 times that standard

Table 6. Ratio of medians drifts and standard errors for the Van Nuys and SAC MDOF systems. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

VN	T1a	T2a	T3a	T1b	T2b	T3b	T	LA9	T1a	T2a	T3a	T1b	T2b	T3b	T
A1a	1.09	1.08	1.00	0.93	1.01	1.01	1.02	A1a	1.04	0.94	0.93	1.07	1.04	1.10	1.02
	0.11	*0.08*	*0.11*	*0.17*	*0.10*	*0.16*	*0.08*		*0.18*	*0.19*	*0.19*	*0.17*	*0.20*	*0.20*	*0.15*
A2a	1.08	1.07	0.99	0.93	1.00	1.00	1.01	A2a	1.18	1.07	1.06	1.22	1.19	1.26	1.16
	0.10	*0.08*	*0.11*	*0.17*	*0.09*	*0.16*	*0.07*		*0.18*	*0.19*	*0.19*	*0.17*	*0.20*	*0.20*	*0.15*
A3a	1.03	1.02	0.95	0.88	0.96	0.96	0.96	A3a	1.04	0.94	0.93	1.07	1.04	1.10	1.02
	0.13	*0.11*	*0.13*	*0.18*	*0.12*	*0.18*	*0.11*		*0.17*	*0.17*	*0.17*	*0.15*	*0.18*	*0.19*	*0.12*
A4a	1.10	1.10	1.02	0.95	1.03	1.03	1.03	A4a	0.93	0.84	0.83	0.95	0.93	0.98	0.91
	0.13	*0.11*	*0.13*	*0.18*	*0.12*	*0.18*	*0.11*		*0.17*	*0.18*	*0.17*	*0.15*	*0.18*	*0.19*	*0.13*
A5a	0.93	0.92	0.85	0.80	0.86	0.86	**0.87**	A5a	0.96	0.88	0.86	0.99	0.97	1.02	0.95
	0.10	*0.07*	*0.10*	*0.16*	*0.08*	*0.16*	*0.06*		*0.17*	*0.17*	*0.16*	*0.15*	*0.18*	*0.18*	*0.12*
A1b	1.16	1.15	1.07	1.00	1.08	1.08	1.09	A1b	0.82	0.74	0.73	0.84	0.82	0.87	0.80
	0.16	*0.15*	*0.17*	*0.21*	*0.16*	*0.21*	*0.15*		*0.18*	*0.19*	*0.18*	*0.16*	*0.19*	*0.20*	*0.14*
A2b	1.05	1.04	0.97	0.90	0.98	0.98	0.99	A2b	1.00	0.91	0.90	1.03	1.01	1.07	0.98
	0.10	*0.08*	*0.11*	*0.16*	*0.09*	*0.16*	*0.07*		*0.22*	*0.23*	*0.23*	*0.21*	*0.24*	*0.24*	*0.20*
A3b	1.00	0.99	0.92	0.86	0.93	0.93	0.93	A3b	1.03	0.93	0.92	1.06	1.03	1.09	1.01
	0.10	*0.08*	*0.11*	*0.16*	*0.09*	*0.16*	*0.07*		*0.16*	*0.17*	*0.16*	*0.14*	*0.18*	*0.18*	*0.12*
A4b	0.93	0.93	0.86	0.80	0.87	0.87	0.88	A4b	0.95	0.86	0.85	0.98	0.95	1.01	0.93
	0.11	*0.09*	*0.11*	*0.17*	*0.10*	*0.17*	*0.08*		*0.15*	*0.16*	*0.15*	*0.13*	*0.16*	*0.17*	*0.10*
A5b	1.13	1.12	1.05	0.97	1.05	1.05	1.06	A5b	0.78	0.71	**0.70**	0.80	0.79	0.83	**0.77**
	0.12	*0.10*	*0.12*	*0.18*	*0.11*	*0.17*	*0.10*		*0.17*	*0.18*	*0.17*	*0.15*	*0.18*	*0.19*	*0.13*
A	1.05	1.04	0.97	0.90	0.97	0.97	0.98	A	0.97	0.88	0.87	1.00	0.97	1.03	0.95
	0.09	*0.06*	*0.09*	*0.16*	*0.07*	*0.15*	*0.05*		*0.13*	*0.14*	*0.13*	*0.10*	*0.15*	*0.15*	*0.06*

error in both examples cited). This observation demonstrates that one must keep this dispersion in mind when making judgments about such matters; the same sample size is much less informative about some systems than others.

Similar conclusions can be found for the 1.5 sigma (15% significance level) basis. Table 7 shows the summary of the number of small sample comparisons showing statistically significant differences for the SDOF cases for both significance levels. These percentages of significant differences should be compared with the fraction of cases (15% and 6%) that one expects to show significant differences even when the hypothesis is true. As discussed above, however, there is some correlation among these 60 ratios, which may affect these expected fractions to some degree. Those for the T=4 sec case are generally high, while the various pooled cases shown are generally lower than expected.

Observation of the typical ratios observed shows that the small sample results may either be above or below unity for all the cases. Looking solely at the large sample ratios, one might want to conclude that the arbitrary records produce somewhat high responses for the trilinear and 0.1-sec bilinear cases, while in the 4-sec bilinear case, they may induce somewhat low responses at higher ductility. But, given the large dispersions,

Table 7. SDOF systems results summary

$\mu \approx 6$	T=1.5 sec	T=0.1 sec	T=4 sec	T=1.5 sec (Tril.)	High μ—All T
1.5	3/60	5/60	14/60	6/60	28/240
2	3/60	2/60	7/60	2/60	14/240
$\mu \approx 2$	T=1.5 sec	T=0.1 sec	T=4 sec	T=1.5 sec (Tril.)	Low μ—All T
1.5	0/60	7/60	13/60	0/60	20/240
2	0/60	0/60	4/60	0/60	4/240
All μ	T=1.5 sec	T=0.1 sec	T=4 sec	T=1.5 sec (Tril.)	All μ—All T
1.5	3/120	12/120	27/120	6/120	48/480
2	3/120	2/120	11/120	2/120	18/480

these conclusions cannot be supported statistically by this data. In fact, taken together, the tables and summaries give little or no support to the notion that one should expect major systematic errors in the estimated nonlinear responses if records are selected simply at random from a catalog with a comparatively wide magnitude and distance range, rather selected carefully, matching the mean or modal {M, R} scenario from a disaggregated PSHA. (It should be pointed out that that scenario is, in any case, based on a disaggregation of simply *linear* SDOF spectral accelerations, not of the nonlinear SDOF or MDOF model of the structure at the moderate to large ductility of interest in such assessments.) While the cases here are admittedly limited and focused on moderate-period systems, the tests do extend to more extreme periods and to nonductile SDOF and MDOF systems. Further, there is other evidence that supports the general conclusion that M and R play at most only minor roles in affecting nonlinear displacements of structures. Studies such as those in Bazzurro and Cornell (1994a and 1994b), Shome et al. (1998), Shome (1999), Carballo (2000), Luco (2002), Medina (2002), and Jalayer (2003) all quantify in various other ways the general insensitivity of nonlinear structural displacements to R and the light sensitivity to M. Cases that may display some sensitivity to magnitude include tall buildings with important second-mode effects and very short period systems. The case of small magnitudes (less than 6, say), which may be important in certain low-seismicity regions and for which the shape of the records Fourier or response spectra may be more strongly magnitude dependent, has not received as much study. It is apparently the magnitude-dependence of the shape of the spectrum that drives any such magnitude dependence.

THE SCALING ISSUE

The validity of the scaling of accelerograms by a scalar intensity measure such as PGA or first-mode spectral acceleration is also one of the objectives that the study above, as it was originally conceived by the writers, would help to address. The concern is usually stated that weaker records scaled up may produce lower responses (or at least different) responses than unscaled records. It was presumed that the arbitrary sets, com-

Table 8. SDOF scaling factors between the arbitrary sets and the original target sets

T=1.5	Scale Factor	T=0.85	Scale Factor	T=0.1	Scalar Factor	T=4	Scalar Factor
A1	1	A1	1.2	A1	1.4	A1	1.1
A2	1.4	A2	1.4	A2	1.4	A2	1.4
A3	1.2	A3	1.0	A3	1.4	A3	1.4
A4	0.9	A4	0.8	A4	1	A4	0.8
A5	1	A5	1.0	A5	0.9	A5	1.3
A	1.1	A	1.1	A	1.2	A	1.2

ing from, on average, smaller magnitudes and larger distances, would have significantly lower strength records, requiring significant scaling to match the target sets. However, as shown in Table 8 the scaling factors for arbitrary-to-target sets (pooled a-b component) are not much larger than one, on average (The last row of Table 8 reports the median of the scaling factors over all the arbitrary sets.) Therefore, the results above cannot serve as a test of this issue. The fact that the scaling factors are not large, in fact, avoids the possibility that any such effect might be confounding the record selection conclusions above.

Not enough records are available in the far-field database to get sufficiently large scaling factors to test the manipulability of the records while still holding fixed the several features discussed above that are the bases of the previous target set choices. Therefore, two new target sets (denoted NT1 and NT2) especially prepared for the scaling issue have been chosen from the PEER catalog still using only records coming from California events with the same instrument and soil characteristics as before, but now with a distance closer than 15 km. Given this short distance range, the concern is whether these "nearby" records can be legitimately treated as if they were representative of non-near-to-the-source or "ordinary" records. They have, however, been selected to avoid, to the degree feasible, records with an obvious pulse-like nature suggesting possible directivity effects. Although in the NT1 target set, one pulse-like record (Northridge 1994/01/17 12:31 77 Rinaldi Receiving Station) is present, the recording used is not the fault-normal component. Further, a check was made, by holding the record out of the set, to confirm that its presence did not introduce any bias in the response results obtained. A statistical test was also made to confirm that the responses from these 2 new target record sets are not statistically different from those of the 6 target sets used above in the selection exercise.

Two of these nearby sets have been prepared. To be sure that these nearby target sets were strong enough to investigate the scaling, one set (denoted NT1) has been chosen by picking the 10 strongest records. The strongest records were defined as those with the maximum *average* spectral acceleration at the four periods (0.1, 0.85, 1.5, and 4 sec). The objective of using this average was to reduce the likelihood of selecting records that happen to be unduly strong at one period due to there being a large peak in its spectrum. Such records would be present if the selection were done structure-by-structure seeking the strongest records based on the spectrum at a single period. Among records scaled to

Table 9. SDOF scaling factors between the arbitrary sets and the new target sets

T=1.5	NT1	NT2	T=0.85	NT1	NT2	T=0.1	NT1	NT2	T=4	NT1	NT2
A1	3.6	1.7	A1	5.2	2.6	A1	4.3	2.9	A1	2.3	2.0
A2	5.5	2.2	A2	6.3	3.1	A2	4.3	2.9	A2	2.8	2.4
A3	4.3	2.0	A3	4.3	2.1	A3	4.2	2.8	A3	2.8	2.4
A4	3.1	1.5	A4	3.5	1.7	A4	3.1	2.0	A4	1.7	1.5
A5	3.6	1.8	A5	4.6	2.3	A5	2.8	1.9	A5	2.6	2.3
A	3.9	1.8	A	4.7	2.3	A	3.7	2.5	A	2.4	2.1

a common single spectral acceleration level, a record with such a peak will generally cause *lower* nonlinear response for structures with that natural period; the common explanation is that as it "softens" it drifts into a regime of lower input power (e.g., Kennedy et al. 1984, Carballo 2000). To the extent such peaks are systematically present in these records, the ratio of arbitrary-to-target medians will be increased. A proxy for the "peakiness" of a record is the normalized residual ("epsilon") between the record Sa value and that predicted by the Abrahamson and Silva (1997) attenuation law. In NT1 they have been found to be on average about 0.5, whereas they are close to zero for the arbitrary sets and for NT2. The target set NT1 can be found in Table A3 (see Appendix).

A second new target set has been built choosing *randomly* among events with the strongest earthquake magnitude in the nearby field arbitrary sets. This set, called NT2, is also listed in Table A3 and shows two records overlapping with NT2. In Table 9 the scaling factors of the same 10 arbitrary sets to the 2 new target sets are shown. These tests were not conducted on the Van Nuys MDOF model. The Van Nuys MDOF has been replaced by a T=0.85 sec SDOF, while the scaling factors for the SAC LA-9 factors can be considered coincident with those for the 1.5-sec SDOF.

The last row of Table 9 reports the median of the scaling factors over all the arbitrary sets. As expected, NT2 set shows lower scaling factors than NT1. The NT2 scaling factors are nonetheless of a useful amplitude (about 2).

SCALING ISSUE RESULTS

The results of the nonlinear analyses and ratio calculations are presented in Tables 10–15. They are in the same format as the tables discussed in the selection study above, except no pooled results are provided across the 2 target sets because the scaling levels are different. The three moderate and short period bilinear SDOF systems show virtually no cases significantly different from one (at the two-sigma or 6%) significance level; that is, they suggest that scaling by factors up to at least 4 does not bias the nonlinear displacement results for these systems at low or high ductility levels. (The 0.85-sec SDOF system standing in for the Van Nuys MDOF case shows some high scaling over prediction at the *smaller* ductility only.) The trilinear strength-degrading system suggests that at the high scaling and high ductility there may be some overestimation of nonlinear displacements by the scaled-up records. But the pooled ratio of 1.66 is not significant sta-

Table 10. Ratio of medians drifts and standard errors for the T = 1.5 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	NT1	NT2	$\mu \approx 2$	NT1	NT2
A1a	1.09	0.95	A1a	0.98	0.88
	0.17	*0.10*		*0.11*	*0.13*
A2a	1.31	1.09	A2a	1.12	1.01
	0.21	*0.15*		*0.12*	*0.14*
A3a	1.16	0.95	A3a	1.03	0.95
	0.19	*0.12*		*0.08*	*0.11*
A4a	1.09	0.89	A4a	1.00	0.91
	0.18	*0.13*		*0.09*	*0.12*
A5a	1.12	0.94	A5a	1.01	0.91
	0.19	*0.13*		*0.11*	*0.13*
A1b	1.25	1.07	A1b	1.07	0.98
	0.18	*0.13*		*0.09*	*0.12*
A2b	1.39	0.98	A2b	1.18	0.99
	0.29	*0.21*		*0.17*	*0.15*
A3b	1.18	0.99	A3b	1.08	0.96
	0.20	*0.14*		*0.08*	*0.12*
A4b	1.21	1.02	A4b	1.10	0.98
	0.19	*0.13*		*0.10*	*0.12*
A5b	0.85	**0.72**	A5b	0.94	0.87
	0.18	*0.12*		*0.10*	*0.12*
A	1.14	0.96	A	1.04	0.95
	0.15	*0.07*		*0.07*	*0.11*

tistically in the face of the very large dispersion. (The log at the median ratio over β is 1.58, just significant at the 15% *level*, and hence underlined). Curiously, at low ductility the arbitrary records are on average *lower* than the target set NT1, and statistically significantly so.

On the same basis, the long-period SDOF system also is significant under high scaling and both high and low ductility, but the scaled records now appear to underestimate the target sets. The exception is at the low ductility and lesser scaling (NT2) when the scaled-up records overestimate the drift to a degree nearly significant at the 15% level. The moderate-period MDOF system shows five rather unexpected and/or contradictory significant cases. The SAC steel frame shows a significantly low ratio, but only for the lower of the two scale-factor target sets (NT2). Recall that it gave only one significant ratio in 60 (Table 6) when the scale factor was close to unity. Therefore, if this NT2 result is meaningful it implies an unlikely nonmonotonic bias with scale factor. Like the other cases above, such "internal inconsistencies" make generalizations difficult. In short, while the scaling results here are limited by the shortage of strong records, there

Table 11. Ratio of medians drifts and standard errors for the T=1.5 sec trilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	NT1	NT2	$\mu \approx 2$	NT1	NT2
A1a	1.15	0.72	A1a	0.61	1.00
	0.67	*0.42*		*0.38*	*0.20*
A2a	**4.21**	1.75	A2a	0.94	1.27
	0.47	*0.53*		*0.13*	*0.25*
A3a	2.21	1.41	A3a	0.86	1.15
	0.47	*0.43*		*0.13*	*0.22*
A4a	1.64	1.00	A4a	0.88	1.02
	0.40	*0.41*		*0.13*	*0.22*
A5a	0.88	1.39	A5a	*0.47*	1.15
	0.79	*0.49*		*0.45*	*0.19*
A1b	1.21	0.92	A1b	0.73	1.06
	0.56	*0.41*		*0.27*	*0.19*
A2b	2.17	1.19	A2b	0.94	1.44
	0.49	*0.51*		*0.14*	*0.31*
A3b	2.22	1.92	A3b	0.89	1.04
	0.47	*0.40*		*0.13*	*0.22*
A4b	**3.09**	1.57	A4b	0.89	1.10
	0.40	*0.45*		*0.13*	*0.22*
A5b	0.50	0.99	A5b	0.39	1.00
	0.78	*0.44*		*0.52*	*0.25*
A	1.66	1.23	A	**0.74**	1.12
	0.32	*0.33*		*0.14*	*0.15*

is no compelling or consistent evidence in this study that scaling of (arbitrarily selected) records causes bias in the estimation of nonlinear displacements of model structures relative to a target set of stronger records.

In general (NT2 is the exception), here the scaling has been from one magnitude-distance bin to another with random selection of records from within that bin. This procedure ensures that there are no systematic peak-valley (or epsilon) effects. Other studies in progress are beginning to show that there can be systematic bias due to scaling if the effects of peaks and valley are not avoided. As described above, records with spectral peaks at the dominant spectral period tend to give smaller nonlinear displacements. The reverse is true for records with spectral valleys. Therefore, if the latter are scaled to the former, overestimation is likely. These effects were minimized by the record selection and processing conducted in this study. Record selection in practice should also avoid unintentional use of record sets dominated by either kind of record.

Table 12. Ratio of medians drifts and standard errors for the T = 0.1 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	NT1	NT2	$\mu \approx 2$	NT1	NT2
A1a	1.24	1.45	A1a	1.19	1.13
	0.37	*0.37*		*0.22*	*0.25*
A2a	0.87	1.00	A2a	1.06	0.99
	0.36	*0.37*		*0.18*	*0.22*
A3a	1.34	1.49	A3a	1.50	1.43
	0.39	*0.39*		*0.29*	*0.31*
A4a	1.19	1.24	A4a	1.35	1.26
	0.40	*0.42*		*0.27*	*0.30*
A5a	0.57	0.64	A5a	0.87	0.83
	0.36	*0.38*		*0.18*	*0.22*
A1b	1.24	1.42	A1b	1.43	1.36
	0.42	*0.42*		*0.29*	*0.31*
A2b	0.83	0.97	A2b	0.97	0.93
	0.36	*0.37*		*0.19*	*0.23*
A3b	1.30	1.57	A3b	1.31	1.24
	0.39	*0.39*		*0.27*	*0.30*
A4b	1.17	1.46	A4b	1.24	1.18
	0.37	*0.36*		*0.24*	*0.26*
A5b	0.65	0.74	A5b	0.90	0.86
	0.29	*0.30*		*0.15*	*0.19*
A	1.00	1.15	A	1.17	1.10
	0.28	*0.27*		*0.15*	*0.19*

SUMMARY OF CONCLUSIONS

Based on the investigation of the nonlinear response of a suite of model structures to sets of records selected to match a specific moderate-magnitude and distance scenario and other moderate-magnitude records selected arbitrarily, this study has found no consistent evidence to suggest that it is necessary to take great care in the selection of records with respect to such factors. The conclusion must be conditional on the characteristics of the uniform catalog available at the time of the study and on the selected magnitude limits. The magnitudes used were limited to moderate values (6.4 to 7.4) because (a) higher values (within the constraints cited above) were not available in the catalog, and (b) smaller values would in practice be unlikely to be chosen for a scenario event in the 7 range, as the catalog does not have an adequate number of 6.4 and larger events and records from which to choose a sample of typical size (10 or less). The mean magnitudes of the A and T sets are 6.6 and 7.1, respectively. The former number suggests, as expected, that the lower magnitudes in the range are more common than the larger. The latter number shows that the T set was indeed selected from the upper tail of the histogram of magnitudes in the catalog. The differential is 0.5 magnitude units. A

Table 13. Ratio of medians drifts and standard errors for the T=4 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	NT1	NT2	$\mu \approx 2$	NT1	NT2
A1a	**0.70**	**0.78**	A1a	**0.77**	0.99
	0.11	*0.09*		*0.10*	*0.06*
A2a	0.89	1.01	A2a	0.92	**1.18**
	0.15	*0.14*		*0.11*	*0.08*
A3a	0.79	0.88	A3a	**0.78**	1.01
	0.12	*0.12*		*0.11*	*0.07*
A4a	0.81	0.91	A4a	0.85	1.08
	0.14	*0.13*		*0.10*	*0.06*
A5a	**0.75**	0.83	A5a	**0.73**	0.93
	0.14	*0.13*		*0.11*	*0.07*
A1b	**0.80**	0.92	A1b	0.90	1.16
	0.11	*0.10*		*0.12*	*0.08*
A2b	0.93	1.04	A2b	0.94	**1.21**
	0.15	*0.14*		*0.11*	*0.07*
A3b	**0.70**	**0.76**	A3b	**0.79**	1.01
	0.11	*0.09*		*0.11*	*0.07*
A4b	**0.72**	0.85	A4b	**0.77**	0.98
	0.14	*0.12*		*0.13*	*0.10*
A5b	0.67	**0.77**	A5b	0.81	1.04
	0.34	*0.13*		*0.32*	*0.08*
A	**0.77**	*0.87*	A	**0.82**	1.06
	0.09	*0.07*		*0.09*	*0.04*

reduction of the lower bound to, say, 6 would have somewhat facilitated meeting the authors' restrictions, designed to avoid overlapping of the samples for the A sets, but it would not have helped the more challenging T set selection. This lower-bound change would have reduced the overlap between records in the T sets with those in the A sets, which would have been somewhat beneficial statistically; it also would have increased the differential in mean magnitudes, which would likely have been a stronger challenge to the posed null hypothesis. As stated, the choice of a lower magnitude of 6.4 was based on the argument that it was the practical choice, while being a full magnitude unit below the largest value.

With respect to distance, a larger catalog, such as that currently under development under PEER, or a larger selected maximum distance (50 kilometers was chosen for reasons analogous to those for the 6.4 lower limit on magnitude) would create a larger mean distance differential between the A and T sets, now 32 versus 25 kilometers. This would not likely cause a greater challenge to the hypothesis because distance per se is known to have little effect on nonlinear response. It would, however, create a larger differential in the mean spectral accelerations, causing a less transparent interaction between the dis-

Table 14. Ratio of medians drifts and standard errors for the T =0.85 sec bilinear SDOF. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

$\mu \approx 6$	NT1	NT2	$\mu \approx 2$	NT1	NT2
A1a	1.01	1.04	A1a	1.08	1.03
	0.17	*0.18*		*0.12*	*0.16*
A2a	1.10	1.14	A2a	1.25	1.20
	0.16	*0.18*		*0.13*	*0.16*
A3a	0.83	0.86	A3a	1.12	1.07
	0.18	*0.19*		*0.09*	*0.13*
A4a	0.82	0.84	A4a	1.15	1.10
	0.16	*0.18*		*0.09*	*0.13*
A5a	0.78	0.81	A5a	0.91	0.87
	0.19	*0.20*		*0.12*	*0.15*
A1b	1.24	1.28	A1b	1.30	1.25
	0.26	*0.27*		*0.14*	*0.17*
A2b	1.36	1.40	A2b	**1.25**	1.20
	0.20	*0.21*		*0.08*	*0.13*
A3b	0.95	0.98	A3b	1.04	1.00
	0.15	*0.17*		*0.09*	*0.13*
A4b	0.91	0.94	A4b	1.05	1.01
	0.18	*0.19*		*0.11*	*0.15*
A5b	1.18	1.22	A5b	**1.39**	1.33
	0.19	*0.20*		*0.13*	*0.16*
A	1.00	1.03	A	1.14	1.10
	0.15	*0.16*		*0.07*	*0.12*

tance and scaling issues. Further, based on the responses of the models to sets of records that are comparatively strong and records that are arbitrarily selected and then scaled up to match the strength of the stronger records, this study has found no compelling evidence that such scaling induces bias in the response estimation. This scaling conclusion reaches to scale factors as high as 4 and ductility up to 6. These general conclusions are limited to firm soil sites and to sites where directivity-induced phenomena are not an important threat. The suite of structures is admittedly limited in number and scope, but it is wide in period range and covers both ductile and strength-degrading backbone models, and both SDOF and MDOF (moment-resisting frame) cases. The scaling study in particular is limited by the number of records. Few qualify as strong enough to test for large scale factors.

Table 15. Ratio of medians drifts and standard errors for the SAC LA-9 MDOF systems. (Bold implies that the hypothesis test of equality is rejected at the 6% significance level.)

	NT1	NT2
A1a	1.02	0.81
	0.15	*0.12*
A2a	1.28	0.90
	0.18	*0.12*
A3a	1.21	0.83
	0.15	*0.11*
A4a	1.09	**0.75**
	0.18	*0.12*
A5a	1.04	**0.79**
	0.16	*0.11*
A1b	0.94	0.67
	0.16	*0.13*
A2b	0.97	0.80
	0.18	*0.18*
A3b	1.17	0.84
	0.17	*0.11*
A4b	1.12	**0.79**
	0.14	*0.10*
A5b	0.88	**0.72**
	0.15	*0.12*
A	1.07	**0.79**
	0.12	*0.08*

ACKNOWLEDGMENTS

The authors would like to thank Dr. David M. Boore for his suggestions and discussions, which improved the significance and accuracy of this paper. Dr. Paolo Bazzurro, Dr. Fatemeh Jalayer, Dr. Nicolas Luco, Dr. Luis Ibarra, and Francesca di Transo have given substantial help and advice for the development of the study. The authors gratefully acknowledge the support provided by the Department of Structural Analysis and Design of the University of Naples to the first author and by the Earthquake Engineering Research Centers Program of the National Science Foundation, under Award Number EEC-9701568 through the Pacific Earthquake Engineering Research Center (PEER), to the second author. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

APPENDIX

Table A1 lists records selected from the PEER database in a "short distance-high magnitude" range to be representative of a specific scenario event that might be the realistic threat at a particular site. Table A2 contains records chosen from a catalog for comparison to the "target sets." These sets were chosen effectively randomly from the catalog without regard to magnitude or distance. Table A3 lists records used to address the scaling issue.

Table A1. Target sets

Set	Event	Station	Record/Component
T1a	Cape Mendocino 1992/04/25 18:06	89509 Eureka—Myrtle & West	CAPEMEND/EUR000
T1a	Cape Mendocino 1992/04/25 18:06	89486 Fortuna—Fortuna Blvd	CAPEMEND/FOR000
T1a	Imperial Valley 1979/10/15 23:16	6622 Compuertas	IMPVALL/H-CMP015
T1a	Imperial Valley 1979/10/15 23:16	6621 Chihuahua	IMPVALL/H-CHI012
T1a	Landers 1992/06/28 11:58	5070 North Palm Springs	LANDERS/NPS000
T1a	Landers 1992/06/28 11:58	12025 Palm Springs Airport	LANDERS/PSA000
T1a	Loma Prieta 1989/10/18 00:05	57382 Gilroy Array #4	LOMAP/G04000
T1a	Loma Prieta 1989/10/18 00:05	14 WAHO	LOMAP/WAH000
T1a	Northridge 1994/01/17 12:31	24389 LA—Century City CC North	NORTHR/CCN360
T1a	Northridge 1994/01/17 12:31	24283 Moorpark—Fire Sta	NORTHR/MRP180
T2a	Cape Mendocino 1992/04/25 18:06	89486 Fortuna—Fortuna Blvd	CAPEMEND/FOR000
T2a	Cape Mendocino 1992/04/25 18:06	89324 Rio Dell Overpass—FF	CAPEMEND/RIO360
T2a	Imperial Valley 1979/10/15 23:16	931 El Centro Array #12	IMPVALL/H-E12140
T2a	Imperial Valley 1979/10/15 23:16	5056 El Centro Array #1	IMPVALL/H-E01140
T2a	Landers 1992/06/28 11:58	5071 Morongo Valley	LANDERS/MVH000
T2a	Landers 1992/06/28 11:58	22074 Yermo Fire Station	LANDERS/YER360
T2a	Loma Prieta 1989/10/18 00:05	1652 Anderson Dam (Downstream)	LOMAP/AND360
T2a	Loma Prieta 1989/10/18 00:05	57066 Agnews State Hospital	LOMAP/AGW000
T2a	Northridge 1994/01/17 12:31	24461 Alhambra—Fremont School	NORTHR/ALH360
T2a	Northridge 1994/01/17 12:31	24055 Leona Valley #5—Ritter	NORTHR/LV5000
T3a	Cape Mendocino 1992/04/25 18:06	89509 Eureka—Myrtle & West	CAPEMEND/EUR000
T3a	Cape Mendocino 1992/04/25 18:06	89324 Rio Dell Overpass—FF	CAPEMEND/RIO360
T3a	Imperial Valley 1979/10/15 23:16	6621 Chihuahua	IMPVALL/H-CHI012
T3a	Imperial Valley 1979/10/15 23:16	5059 El Centro Array #13	IMPVALL/H-E13140
T3a	Landers 1992/06/28 11:58	23559 Barstow	LANDERS/BRS000
T3a	Landers 1992/06/28 11:58	12149 Desert Hot Springs	LANDERS/DSP000
T3a	Loma Prieta 1989/10/18 00:05	57504 Coyote Lake Dam (Downst)	LOMAP/CLD195
T3a	Loma Prieta 1989/10/18 00:05	1695 Sunnyvale—Colton Ave.	LOMAP/SVL360
T3a	Northridge 1994/01/17 12:31	24303 LA—Hollywood Stor FF	NORTHR/HOL360
T3a	Northridge 1994/01/17 12:31	24309 Leona Valley #6	NORTHR/LV6360

Table A2. Arbitrary sets

Set	Event	Station	Record/Component
A1a	Cape Mendocino 1992/04/25 18:06	89509 Eureka—Myrtle & West	CAPEMEND/EUR000
A1a	Coalinga 1983/05/02 23:42	36410 Parkfield—Cholame 3W	COALINGA/H-C03000
A1a	Coalinga 1983/05/02 23:42	36444 Parkfield—Fault Zone 10	COALINGA/H-Z10000
A1a	Imperial Valley 1979/10/15 23:16	5059 El Centro Array #13	IMPVALL/H-E13140
A1a	Imperial Valley 1979/10/15 23:16	5052 Plaster City	IMPVALL/H-PLS045
A1a	Landers 1992/06/28 11:58	22074 Yermo Fire Station	LANDERS/YER360
A1a	Loma Prieta 1989/10/18 00:05	47524 Hollister—South & Pine	LOMAP/HSP000
A1a	Northridge 1994/01/17 12:31	14368 Downey—Co Maint Bldg	NORTHR/DWN360
A1a	San Fernando 1971/02/09 14:00	125 Lake Hughes #1	SFERN/L01021
A1a	Superstitn Hills(B) 1987/11/24 13:16	5210 Wildlife Liquef. Array	SUPERST/B-IVW360
A2a	Cape Mendocino 1992/04/25 18:06	89324 Rio Dell Overpass—FF	CAPEMEND/RIO270
A2a	Coalinga 1983/05/02 23:42	36408 Parkfield—Fault Zone 3	COALINGA/H-COH000
A2a	Coalinga 1983/05/02 23:42	36439 Parkfield—Gold Hill 3E	COALINGA/H-GH3000
A2a	Imperial Valley 1979/10/15 23:16	724 Niland Fire Station	IMPVALL/H-NIL360
A2a	Imperial Valley 1979/10/15 23:16	5169 Westmorland Fire Sta	IMPVALL/H-WSM180
A2a	Landers 1992/06/28 11:58	12025 Palm Springs Airport	LANDERS/PSA000
A2a	Loma Prieta 1989/10/18 00:05	47179 Salinas—John & Work	LOMAP/SJW160
A2a	Northridge 1994/01/17 12:31	24271 Lake Hughes #1	NORTHR/LH1000
A2a	San Fernando 1971/02/09 14:00	262 Palmdale Fire Station	SFERN/PDL120
A2a	Superstitn Hills(A) 1987/11/24 05:14	5210 Wildlife Liquef. Array	SUPERST/A-IVW360
A3a	Cape Mendocino 1992/04/25 18:06	89486 Fortuna—Fortuna Blvd	CAPEMEND/FOR000
A3a	Coalinga 1983/05/02 23:42	36455 Parkfield—Vineyard Cyn 1E	COALINGA/H-PV1000
A3a	Coalinga 1983/05/02 23:42	36447 Parkfield—Vineyard Cyn 2W	COALINGA/H-VC2000
A3a	Imperial Valley 1979/10/15 23:16	5066 Coachella Canal #4	IMPVALL/H-CC4045
A3a	Imperial Valley 1979/10/15 23:16	6605 Delta	IMPVALL/H-DLT262
A3a	Landers 1992/06/28 11:58	12149 Desert Hot Springs	LANDERS/DSP000
A3a	Loma Prieta 1989/10/18 00:05	1002 APEEL 2—Redwood City	LOMAP/A02043
A3a	Loma Prieta 1989/10/18 00:05	57191 Halls Valley	LOMAP/HVR000
A3a	Northridge 1994/01/17 12:31	14403 LA—116th St School	NORTHR/116360
A3a	San Fernando 1971/02/09 14:00	289 Whittier Narrows Dam	SFERN/WND143
A4a	Coalinga 1983/05/02 23:42	36407 Parkfield—Fault Zone 1	COALINGA/H-COW000
A4a	Coalinga 1983/05/02 23:42	36456 Parkfield—Fault Zone 14	COALINGA/H-Z14000
A4a	Imperial Valley 1979/10/15 23:16	6621 Chihuahua	IMPVALL/H-CHI012
A4a	Imperial Valley 1979/10/15 23:16	931 El Centro Array #12	IMPVALL/H-E12140
A4a	Landers 1992/06/28 11:58	23559 Barstow	LANDERS/BRS000
A4a	Loma Prieta 1989/10/18 00:05	1652 Anderson Dam (Downst)	LOMAP/AND360
A4a	Loma Prieta 1989/10/18 00:05	1656 Hollister Diff. Array	LOMAP/HDA165
A4a	Northridge 1994/01/17 12:31	14196 Inglewood—Union Oil	NORTHR/ING000
A4a	Northridge 1994/01/17 12:31	24283 Moorpark—Fire Sta	NORTHR/MRP180
A4a	San Fernando 1971/02/09 14:00	135 LA—Hollywood Stor Lot	SFERN/PEL180
A5a	Coalinga 1983/05/02 23:42	36227 Parkfield—Cholame 5W	COALINGA/H-C05360
A5a	Coalinga 1983/05/02 23:42	36441 Parkfield—Vineyard Cyn 6W	COALINGA/H-VC6000
A5a	Imperial Valley 1979/10/15 23:16	5056 El Centro Array #1	IMPVALL/H-E01140
A5a	Imperial Valley 1979/10/15 23:16	6617 Cucapah	IMPVALL/H-QKP085
A5a	Landers 1992/06/28 11:58	5071 Morongo Valley	LANDERS/MVH000
A5a	Landers 1992/06/28 11:58	5070 North Palm Springs	LANDERS/NPS000
A5a	Loma Prieta 1989/10/18 00:05	57504 Coyote Lake Dam (Downst)	LOMAP/CLD195
A5a	Loma Prieta 1989/10/18 00:05	57382 Gilroy Array #4	LOMAP/G04000
A5a	Northridge 1994/01/17 12:31	24157 LA—Baldwin Hills	NORTHR/BLD360
A5a	Northridge 1994/01/17 12:31	24303 LA—Hollywood Stor FF	NORTHR/HOL360

Table A3. Scaling target sets

Set	Event	Station	Record/Component
NT1	Coalinga 1983/05/02 23:42	1162 Pleasant Valley P.P.—yard	COALINGA/H-PVY045
NT1	Coalinga 1983/05/02 23:42	1162 Pleasant Valley P.P.—yard	COALINGA/H-PVY135
NT1	Imperial Valley 1979/10/15 23:16	5054 Bonds Corner	IMPVALL/H-BCR230
NT1	Imperial Valley 1979/10/15 23:16	5028 El Centro Array #7	IMPVALL/H-E07230
NT1	Loma Prieta 1989/10/18 00:05	47125 Capitola	LOMAP/CAP000
NT1	Loma Prieta 1989/10/18 00:05	57007 Corralitos	LOMAP/CLS090
NT1	Northridge 1994/01/17 12:31	77 Rinaldi Receiving Sta	NORTHR/RRS228
NT1	Northridge 1994/01/17 12:31	74 Sylmar—Converter Sta	NORTHR/SCS142
NT1	N. Palm Springs 1986/07/08 09:20	5070 North Palm Springs	PALMSPR/NPS210
NT1	Superstitn Hills(B) 1987/11/24 13:16	5051 Parachute Test Site	SUPERST/B-PTS225
NT2	Imperial Valley 1979/10/15 23:16	5054 Bonds Corner	IMPVALL/H-BCR140
NT2	Imperial Valley 1979/10/15 23:16	5060 Brawley Airport	IMPVALL/H-BRA315
NT2	Imperial Valley 1940/05/19 04:37	117 El Centro Array #9	IMPVALL/I-ELC180
NT2	Landers 1992/06/28 11:58	22170 Joshua Tree	LANDERS/JOS000
NT2	Loma Prieta 1989/10/18 00:05	47381 Gilroy Array #3	LOMAP/G03000
NT2	Loma Prieta 1989/10/18 00:05	47006 Gilroy—Gavilan Coll.	LOMAP/GIL337
NT2	Northridge 1994/01/17 12:31	90053 Canoga Pk—Topanga Cyn	NORTHR/CNP196
NT2	Northridge 1994/01/17 12:31	90009 N. Hollywood—Coldwater Cyn	NORTHR/CWC180
NT2	Superstitn Hills(B) 1987/11/24 13:16	01335 El Centro Imp. Co. Cent	SUPERST/B-ICC000
NT2	Superstitn Hills(B) 1987/11/24 13:16	5051 Parachute Test Site	SUPERST/B-PTS225

REFERENCES

Abrahamson, N. A., and Silva, W. J., 1997. Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismol. Res. Lett.* **68** (1), 94–127.

Bazzurro, P., and Cornell, C. A., 1994a, Seismic hazard analysis of nonlinear structures. I: Methodology, *J. Struct. Eng.* **120**, 3325–3344.

Bazzurro, P., and Cornell, C. A., 1994b, Seismic hazard analysis of nonlinear structures. II: Applications, *J. Struct. Eng.* **120**, 3345-3365.

Benjamin, J. R., and Cornell, C. A., 1970. *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, London.

Carballo Arévalo, J. E., 2000. Probabilistic Seismic Demand Analysis: Spectrum Matching and Design, Ph.D. thesis, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA. (http://www.stanford.edu/group/rms/)

Foutch, D. A., and Shi, S., 2002. Modeling of steel moment frames for seismic loads, *J. Constr. Steel Res.* **58** (5–8).

Gupta, A., and Krawinkler, H., 1999. Seismic Demands for Performance Evaluation of Steel Moment Resisting Frame Structures (SAC Task 5.4.3), *Report No. 132*, John A. Blume Earthquake Engineering Center, Stanford University, Stanford, CA.

Iervolino, I., 2004. Record Selection for Nonlinear Seismic Analysis of Structures, M.Sc. thesis, ROSE School, Institute for Advanced Study, University of Pavia, Italy. (http://www.roseschool.it/files/research-students-MD2004-iervolino.htm)

International Code Council, Inc. (ICC), 2000. *International Building Code*, Falls Church, VA.

Jalayer, F., 2003. Direct Probabilistic Seismic Analysis: Implementing Non-Linear Dynamic Assessments, Ph.D. thesis, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA. (http://www.stanford.edu/group/rms/)

Kennedy, R. P., Short, S. A., Mertz, K. L., Tokarz, F. J., Idriss, I. M., Power, M. S., and Sadigh, K., 1984. *Engineering Characterization of Ground Motion—Task I: Effects of Characteristics of Free-Field Motion on Structural Response*, NUREG/CR-3805 **1**, U.S. Nuclear Regulatory Commission.

Luco, N., 2002. Probabilistic Seismic Demand Analysis, SMRF Connection Fractures, and Near-Source Effect, Ph.D. thesis, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA. (http://www.stanford.edu/group/rms/)

Luco, N., and Cornell, C. A., 2000. Effects of connection fractures on SMRF seismic drift demands, *J. Struct. Eng.* **126** (1).

McGuire, R. K., 1995. Probabilistic seismic hazard analysis and design earthquakes: Closing the loop, *Bull. Seismol. Soc. Am.* **85** (5).

Medina, R. A., 2002. Seismic Demands for Nondeteriorating Frame Structures and their Dependence on Ground Motion. Ph.D. thesis, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA.

Porter, K. A., Beck, J. L., and Shaikhutdinov, R. V., 2002. *Investigation of Sensitivity of Building Loss Estimates to Major Uncertain Variables for the Van Nuys Testbed*, PEER Report Series **2002/03**, August.

Shome, N., 1999. Probabilistic Seismic Demand Analysis of Nonlinear Structures, Ph.D. thesis, Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA. (http://www.stanford.edu/group/rms/)

Shome, N., Cornell, C. A., Bazzurro, P., and Carballo, J. E., 1998. Earthquakes, records and nonlinear responses, *Earthquake Spectra* **14** (3), 469–500.

U.S. Nuclear Regulatory Commission, 2001. *Technical Basis for Revision of Regulatory Guidance on Design Ground Motions: Hazard- and Risk-Consistent Ground Motion Spectra Guidelines, NUREG/CR-6728*, Government Printing Office, Washington, D.C.

Van Nuys Testbed Committee, 2002. PEER Performance-Based Earthquake Engineering Methodology: Structural and Architectural Aspects, Pacific Earthquake Engineering Research Center University of California, Berkeley. (http://www.peertestbeds.net)

(Received 6 May 2004; accepted 16 November 2004)