Prof. C. Allin Cornell
Professor Emeritus
Stanford University
110 Coquito Way
Portola Valley, Ca, US, 94028
cornell@stanford.edu

Probabilistic methods in structural engineering

Luis Esteva has been one of my two most knowledgeable and influential tutors in earthquake engineering and probability. In the summer of 1966 Luis and I were both working at UNAM beside my second such tutor, Emilio Rosenblueth. But Luis was the one who was then also deeply interested in the questions of earthquake occurrence and ground motion prediction, and their representation for engineering hazard assessment purposes. My own work in that field owes a great deal to the openness and patience with which Luis discussed his ongoing studies. His example has also set a standard for me with respect to the importance of keeping our sometimes seemingly esoteric work in probabilistic methods grounded in the needs of engineering practice. The two of us have since shared many concentrated technical and pleasant social hours not only at UNAM and at MIT and Stanford, where I sought his multi-month visits, but also in the many cities around the world that this profession takes you. For all that, Luis Esteva, I thank you and salute you on this special day in your career.

Allin Cornell,
August 10, 2005

ON EARTHQUAKE RECORD SELECTION FOR NONLINEAR DYNAMIC ANALYSIS

ABSTRACT

The apparently simple act of selecting accelerograms for use in conducting nonlinear dynamic analysis has been the subject recent study (e.g., Baker 2005c) and conclusions are emerging. This paper attempts to elucidate those conclusions by starting from very ideal and simplified cases to deduce observations that are supported by more through numerical studies of nonlinear MDOF structural models. A direct and an indirect approach are discussed. To insure both lack of bias and adequate confidence limits in the estimation of response and response likelihoods, the selection of the spectral shape and number of the records is found to be dependent on the site, the structure and ground motion level. With proper choice of the spectral shape scaling of the records should not cause significant bias.

Introduction

It might be said that both Luis Esteva's career and my own have been highly focused on a single problem: estimating the annual frequency, λ, that an earthquake induces in a particular structure some specified behavior state. We have worked separately, together and with many colleagues and students on pieces of and the whole of this issue. There is a half century of literature on this subject with many approaches, simplifications, and results. Yet many of us are still working hard on the subject. Modern computational resources have accelerated the progress, opening opportunities, for example, for large statistical samples of nonlinear dynamic analyses of multi-degree-of-freedom models of structures both for research and for applications. But the increased computational power is also being used by the structural modelers to improve the detail and accuracy of their FEM models. This verisimilitude is especially important as we attempt to push our studies to the extreme damage and collapse domains of behavior important to life safety assessment. Therefore there will always be a demand to limit the number of dynamic analyses used to estimate λ. Even when, as in research contexts, computational limits may be less restrictive than in practice, there remain open many questions about which accelerograms to run in any particular situation. These questions arise both in practice and research.

Current best record-selection practice even in research (e.g., ASCE 2005, Stewart 2002) has the seismologist providing of the order of 10 or fewer records that represent magnitudes and distances identified by probabilistic seismic hazard analysis (PSHA) disaggregation (e.g., Bazzurro 1999) to be the most likely to have caused the event that a particular response spectral ordinate equals the level associated with a particular mean annual rate of exceedance. These records are then scaled as necessary to match the level of the uniform hazard spectrum in one manner or another. An important implication of this practice is that the result is dependent on the character of the seismicity that surrounds the site and the mean return period of interest. It is implicit too that it believed that the magnitude and distance of a record do or may affect the structural response. It is also clear that nothing beyond single-degree-of-freedom (SDOF) linear structural behavior is used in the selection. While many questions surround various elements of this practice (e.g., the number of records, the impact of the scaling, etc.), especially for application to rare, severe response of nonlinear multi-degree-of-freedom (MDOF) structures, it has only recently begun to be investigated for accuracy and effectiveness.

The objective of this paper is to back up briefly to look at the problem from a simple but perhaps fundamental perspective that may give us insights into what to look for and what to avoid in selecting accelerograms for nonlinear time history analysis. This may also suggest some of the compromises we are making in current practice and research.

With these objectives I shall be making, for the purpose of clarity of exposition, various specializations and restrictions and simplifications, most of which could be expanded or generalized without significant effort. For example I take it as a given that the fundamental objective is the estimation of λ. Note that special cases include, for example, the annual frequency (or, approximately, probability) of collapse, of roof drift angle greater than 3%, of maximum interstory drift in the first floor greater than its (random) capacity, etc. Such limit state frequencies are at the basis modern approaches to earthquake codes, guidelines, advanced practice, and performance-based earthquake engineering. They are the first step toward estimation of more direct decision variables involving consequences such as lives lost, economic damage to structure or non-structural elements, and lost occupancy time. Other limitations in the paper will be simple site hazard and structural representations.

Direct Approaches

In the best of circumstances λ would be estimated by monitoring the response of the building itself for a sufficient number of years to estimate λ with the desired accuracy. Suppose for simplicity that single response variable, θ, is of concern (say the maximum interstory drift ratio - MIDR - in a frame). Then, just as with empirical flood frequency analysis, we would need only order and plot these values versus i/n (where n is the number of years of observation) as an empirical complementary cumulative mean annual rate of exceedance function (CCDF), $\lambda_\theta(x)$. The value of x at $\lambda_\theta(x) = 0.001$ would be the estimated 1000 year mean return period MIDR. Even, however, if the structure had been built and monitored, for safety level λ_θ's of the order of 10^{-3} or less, this would require 10,000 or more years of data[1].

Somewhat more realistically the response will be estimated from linear/nonlinear dynamic analysis of a numerical model of the structure, be it an existing building or a proposed design. (We assume here that such numerical models are precise.) Then we would need "only" to have had an accelerometer in place at the site for these 10,000 years. All the records (above some

---

[1] The basis for this statement is that simple binomial trials statistics suggests it requires a sample size of about 10/p to estimate small probability p with a standard error of about +/- 30% of p. To reduce this number to 15% would require 4 times as many years.

practical threshold, say 1000 in total) would need to be run and their resulting values of θ plotted as above. While this case remains completely unrealistic it is an excellent hypothetical model to repeatedly return to as we consider how to address practical record selection. This record set of some 1000 records would implicitly contain records generated by a multitude of magnitudes from different sources at various distances and azimuths over various travel paths, and this set would fully and properly represent all the characteristics of ground motion that might possibly affect the structure's response. What is more, all these characteristics, including site effects, all appear in this large hypothetical set of *in situ* recordings in exactly the correct relative frequency of occurrence, both marginally and jointly. Our objective in practical record selection is to reproduce this condition or, rather, to approximate it as best we can.

Lacking this ideal set of records at our site we turn either to the catalogue of recordings or to simulation. Not wishing in this short note to take on the questioning of just how realistic various current modes of simulating accelerograms are, I limit myself to the catalogue[2]. I further presume that all these recordings are free from instrumentation limitations. Many good catalogues (or virtual catalogues) are readily available today with thousands of records (e.g., PEER 2005). Going to the catalogue of accelerograms recorded elsewhere to represent what has happened at my single site over years is an example of "trading space for time". It carries with it the need to try to understand what is important to the problem at hand to gain confidence that the trade has been fair and the conclusions accurate. So we must ask which records in the catalogue with which characteristics are "right" for my site and in what proportion should select them.

Suppose first, again for simplicity of exposition, that the threat at our site is a single fault segment, located R kilometers from the site[3], which produces only "characteristic" events, i.e., full segment ruptures with very similar magnitudes[4], M, and which does so in a Poisson way with known[5] mean rate λ_o. Then, hoping/assuming (again for simplicity, but close to current practice) that M and R are sufficient representations of the source and path and that, say, "firm soil" (or some like category among the current catalogue representations) is a sufficient characterization of our site, we might logically sort the catalog for all such records and run them through the numerical analysis to obtain values of θ. Ordering and plotting them versus i/m (where m is the number of records found) would produce what we would hope to be a reasonable and accurate estimate of $G_\theta(x)$, the complementary cumulative distribution function (CCDF) of θ given an event on the fault. With this result it follows that $\lambda_\theta(x) = \lambda' \, P[\theta>x|\text{ event}] = \lambda' \, G_\theta(x)$. How much data does this require? Suppose, as in coastal California, λ' is 1/ (several hundred years), then for $\lambda_\theta(x)$ of the order of 10^{-3} to 10^{-4} we need $G_\theta(x)$ to be of the order of 0.1. Reasonably confident estimation requires m to be about[6] 100. Apart from the computational cost, there, of course, are not nearly enough such specific (M, R, firm soil) records in current or foreseeable catalogues (especially as one would like them to be from 100

[2] For approaches using simulated records see, for example, Han 1997, Luco 2002 or Jalayer 2005.
[3] Measured in some relevant way, e.g., as closest distance to the rupture surface.
[4] A more fundamental approach might be to use rupture length rather than the more heavily processed M to characterize events and select records.
[5] Again we assume that this is known accurately. Indeed we shall assume throughout this and all such seismicity information to follow is perfectly known.
[6] Making "a distribution assumption", i.e., fitting this data to a named probability law such as one with an exponential or power form, and estimating its parameter values can reduce this number by a factor of 2 to 3 at the risk of inducing errors in the upper tail by having made the wrong parametric modeling assumption.

independent events). Therefore we must relax the constraints and accept events within an interval[7] M+/-Δ_M and R+/- Δ_R. Immediate questions are: how wide need these bins be to gather an adequate sample size? How much accuracy in θ is given up by using the "wrong" M and R, and how does that increase as say Δ_M gets larger? And are there ways to modify the records to make them "more nearly right"? The simplest, common illustration of this is, say, scaling up by some amount the accelerograms that are from "too small" M's or "too large" R's. And does such scaling induce biases of it own? Before addressing such questions, let us release at least one of the previous unrealistic limitations.

We need to recognize that the assumption of a simple single {M, R} scenario is not only unrealistic but impacts our record selection discussion. Even with a single dominant neighboring fault there may be lesser magnitudes at various locations (and distances) that contribute to the likelihood of exceeding any θ level, x. Further there are inevitably other seismic sources that also contribute to the hazard. These additional scenarios complicate record selection to the degree that they needed to be represented in the analysis. For such cases we need write:

$$\lambda_\theta(x) = \sum_i G_{\theta|M,R}(x \mid m_i, r_i)\, \lambda(m_i, r_i) \qquad (1)$$

in which the individual mean rates of occurrence, λ, and conditional CCDF's, G, are identified for each of the interesting (here discretized) set of {M,R} scenarios. Records would have to be selected as above for each such scenario. Even if there are only 5 to 10 such scenarios the number of records and analyses needed may mount once again into the 1000 range.

Intensity Measure-based Approaches: In Principle

To address these challenges the seismic community has used for decades the notion of what some of us now call an "intensity measure" (IM); examples include PGA and first-mode-period spectral acceleration, $S_a(T_1)$, to serve as an intermediate variable in such assessments. The estimation of the mean rate of exceedance of the IM is the subject of PSHA and the responsibility of earth scientists not structural analysts. The equation for $\lambda_{IM}(y)$ looks just like Eq. 1 with $G_{\theta|M,R}$ replaced by $G_{IM|M,R}$. The last function is obtained from standard strong ground motion attenuation "laws". This PSHA problem is again outside the scope of this discussion; we assume that $\lambda_{IM}(y)$ is available for almost any IM we might chose. It is important to recognize that this a site-specific product reflecting all the {M,R} scenarios. Therefore it has captured a major portion of the specific nature of our site. In particular it has released us from the need to have 10,000 years of recordings at the site in order to measure how frequently important magnitudes occur at critical distances from the site, and how certain measures (IM'S) of the amplitudes of their motions depend on magnitude and attenuate with distance. How completely and adequately this job has been completed for the objective of structure-specific $\lambda_\theta(x)$ estimation remains to be discussed.

[7] Let us drop further explicit discussion of soil conditions. This is not to say that this is not a very important issue. Some portion of the variability in records from similar but different sites would presumably not be found from event to event at the same site. This question is related to the current discussion ergodicity (Anderson 1999)

Returning to the estimation of $\lambda_{\theta}(x)$, it is clear that IM's can likely play a key role. Let us see how. The total probability theorem (e.g., Benjamin 1970) always permits an expansion of $\lambda_{\theta}(x)$ into (e.g., Bazzurro 1998).

$$\lambda_{\theta}(x) = \int G_{\theta|IM}(x \mid y) | \, d\lambda_{IM}(y) | \tag{2}$$

in which the last factor is the absolute value of the derivative of $\lambda_{IM}(y)$, or, loosely, the mean rate of occurrence of a value of the IM equal to y. Having, as we do, $\lambda_{IM}(y)$, our problem is transformed into estimating $G_{\theta|IM,}$ the conditional probability that $\theta > x$ given IM = y. Consider first our original ideal case where we have some 10,000 years of recordings directly from our site. For a given level of y, we would select a random sample from those records that have this (or approximately) this IM level, analyze the structure under them, and process as usual to estimate $G_{\theta|IM}$ versus x for this y level.

How many records would this require? Suppose we are interested again in about 10^{-3} to 10^{-4}. Then we should look at levels of y such that $\lambda_{IM}(y)$ is in this range as well, since it is well known that $\lambda_{IM}(y)$ not $G_{\theta|IM}$ dominates the integral in Eq. 2. Indeed this is so true that generally[8] the estimation of $G_{\theta|IM}$ can be reduced to estimating well the *mean* of θ and only roughly its standard deviation and distribution shape. Accepting this statement as at least a rough approximation we can estimate the required sample size to be of order[9] only 10. This enormous reduction in the necessary sample size is possible because the variability in θ given IM is small relative that in θ marginally. This reduction is the consequence of the PSHA analysis of the IM "absorbing" most of the total variability in θ, leaving only the conditional variability of θ given IM to be estimated from the structural dynamic analyses. Under some practical circumstances only one such level (or "stripe") may be adequate, in particular if there is interest in only a single level of x (e.g., 3% drift) or of mean frequency (e.g., 2% in 50 years); in this case this an extremely attractive approach to the record selector and structural analyst. Unlike the previous approach this number is independent of the number of sources contributing to the hazard. However more than one level of y may be necessary; 3 to 5 or more levels may be required if $\lambda_{\theta}(x)$ is needed for a broad range of x values[10] (Jalayer 2003), raising the required sample size to order 100. In any case this IM-based approach is in principle very efficient[11] and accurate, at least in the ideal case when the large site-specific sample is available. Even with this nominal perfect sample there remain interesting questions as to what variable is best to use as the IM. For MIDR prediction, PGA for example leads to larger required sample sizes than $S_a(T_1)$ simply because it is less well correlated with MIDR for moderate to long period structures. We shall not pursue this question of which IM to use here except (later) to the degree it impacts our focus: record selection.

---

[8] Exceptions include case where the IM hazard curve is very steep and the dispersion of θ given IM is large in the region of interest.

[9] The basis for this is that the log of $\lambda_{\theta}(x)$ is roughly proportional 2 to 3 times the mean of the log of x. Therefore the +/-30% standard error in λ_{θ} mentioned above requires about a 10% standard error in the mean of x. This in turn requires a sample size equal to the square of the coefficient of variation (COV) of x divided by this 0.1. The COV of say MIDR of an MDOF frame structure given a reasonable choice of IM (such as Sa(T1)) is less than 0.3 to 0.5 plus for very severe degrees of nonlinearity.

[10] Especially near global collapse when the dependence on y may be very nonlinear and the dispersion broad.

[11] It can be interpreted as an application of conditional Monte Carlo analysis.

Intensity Measure-based Approaches: In Practice

Now we must return to the real world and ask how records are to be selected in the IM-base approach when we must depend not on a ideal site-specific catalog but on the existing catalog of recordings. This question is commonly asked and answered in practice but rarely in a very formal way. Given the discussion above it should not be a surprise that this is not a trivial formal question. There is no error in the application in the formulation of Eq. 2 and we have assumed that we have an accurate site-specific assessment of $\lambda_{IM}(y)$. Therefore the problem reduces to how one should select records to estimate $G_{\theta|IM}$. The simple answer is we chose them so that we get the right answer. This is more complex than simply what is an adequate sample size because there is now the question of suitability of records recorded elsewhere to this site. The contention is that this question can be addressed only if one considers carefully the structure of concern (and the IM at hand). Let us consider two simple structures to get a sense of what is involved.

Single DOF Linear Structure

Suppose to begin that our structure is simply a linear oscillator, and that we have selected the IM to be the common one: $S_a(T_1)$ where T1 is the natural period of the structure. Our objective is to select records from an available catalog to estimate $G_{\theta|IM}(x|y)$ accurately. In this case record selection for dynamic analysis purposes is clearly trivial. One can select *any* record (no matter, for example, what its M, R, or relative strength), scale it such that its IM = y, run the dynamic analysis and get, of course[12], $\theta = y$ because in this case IM and θ are the same. Further as all records will produce the same value, only a single record is necessary for perfect accuracy. In addition the same record can be used perfectly accurately for all levels, y, of the IM, only scaling is required. This absolute robustness with respect to record selection and scaling and this low dispersion (with its implied computational efficiency) are all a direct result of the selection of the IM, and the simple structure. Had the IM been another popular choice, PGA, none of these conclusions would hold. Because many structures are known to be "first-mode-dominant" it follows further that these desirable properties of $S_a(T_1)$ as an IM are likely to hold to some degree for many real structures, at least in the linear range. Therefore, it can be anticipated that many concerns about record selection for such cases are unwarranted. Dynamic analysis of linear first-mode-dominated structures is seldom needed, however. Nonetheless it suggests that this IM is a natural starting point from which to look for improvements for more complex structures. Certainly those candidates for IM, such as PGA, that create record selection sensitivity for this simplest structural problem are not strong initial choices for the more realistic structural problems.

Two DOF Linear Structure

Suppose next that our structure is, say, a two-story frame that can be represented accurately by a

[12] Provided, as I assume here for argument's sake, our structure has 5% damping as is standard for the attenuation laws used in PSHA. If our structure has a different damping level there will be an offset dependent on average on the degree to which the two damping levels differ and there will be some comparatively mild dispersion from record to record. There is no evidence that this effect is dependent on M or R or other such record properties.

2-DOF linear model. Further continue to assume that the IM is still $S_a(T_1)$, where T_1 is the first-mode period. How now should the catalog be searched for records for dynamic analyses to be used? For the purposes of illustrating the principles, let us further assume that in fact the simple square-root-of-sums-of-squares (SRSS) approximation is exact[13]. In this case given that IM = $S_a(T_1) = y$, the response of interest can be written:

$$\theta = \sqrt{c_1^2\, y^2 + c_2^2\, S_a^2(T_2)} \tag{3}$$

in which c_1 and c_2 are coefficients depending on the dynamic properties of the structure. This form makes it clear that (given $S_a(T_1)$) θ is simply a function of the random variable $S_a(T_2)$, where the probability distribution of $S_a(T_2)$ is the *conditional* distribution of $S_a(T_2)$ given $S_a(T_1)$. We shall take advantage of this below. One can also re-write Eq. 3 as

$$\theta = c_1 y \sqrt{1 + \frac{c_2^2\, S_a^2(T_2)}{c_1^2\, y^2}} \tag{4}$$

In this form it is clear, first, that our concern is only with cases in which the second mode makes a comparatively strong contribution to θ, as indicated by the ratio under the radical, and, second, that θ is simply a function of the (random) spectral ratio $R_{2/1} = S_a(T_2) / S_a(T_1)$ – still conditioned, of course, on $S_a(T_1) = y$. The former observation implies that the record selection is trivial and robust in the first-mode dominated case, as discussed above. The latter observation emphasizes that once the IM level is given it is the only relative value, $S_a(T_2)/y$, (or *spectral shape*) that matters. This notion carries over to other structures as well. Knowledge of this fact supports the practice of selecting records from {M, R} bins that dominate the site hazard, because M is known to have some effect (in the mean at least) on spectral shape. The fact is even more evident in the common practice of first selecting records whose spectral shapes closely match that of the UHS or of the median spectrum given the dominant M and R, and then scaling them to match the level of the target spectrum. Neither of these practices, however, reflects the *conditional* nature of this dependence. We address this issue this next.

To estimate well $G_{\theta|IM}(x|y)$ when the second mode is important we clearly need to select records from the catalog that capture accurately the conditional distribution of $S_a(T_2)$ given $S_a(T_1) = y$ (or equivalently of $R_{2/1}$ given $S_a(T_1) = y$). This distribution is not readily available today[14]. But we can get guidance as follows.

In order to understand better how in principle record selection should proceed in this case, we shall once again simplify by assuming, for the moment, that a single M = m and R = r scenario dominates our site's hazard. Conventional attenuation laws are based on the fact any Sa(T) value has a lognormal distribution with the mean of the natural log, call it $\mu_{\ln S}$, equal to specified function of m and r and with standard deviation of the log, $\sigma_{\ln S}$, equal to another such function. It is reasonable to assume further that the spectral accelerations at two different periods are jointly

[13] Again one might ask: then why is dynamic analysis necessary? It is not, of course, but I again appeal to the sake of the argument.

[14] It happens that they can be found from disaggregation of vector-valued PSHA (Bazzurro 2002), public tools for which are under development (Somerville 2005).

lognormal with correlation coefficient[15] of the logs ρ. In this case the conditional mean of log $S_a(T_2)$ given $S_a(T_1) = y$ is:

$$E[\ln(S_a(T_2)|\,S_1 = y] = \mu_{\ln S_a(T_2)} + \rho \frac{\sigma_{\ln S_a(T_2)}}{\sigma_{\ln S_a(T_1)}}(\ln y - \mu_{\ln S_a(T_1)}) \tag{5}$$

The median, η, of $S_a(T_2)$ given $S_a(T_1) = y$ is *e* raised to this power or:

$$\eta_{S_a(T_2)|S_1(T_1)=y} = \eta_{S_a(T_2)} \exp[\rho\sigma_{\ln S_a(T_2)}\, \varepsilon_{S_a(T_1)}\,] \tag{6}$$

in which $\varepsilon_{S_a(T_1)} = (\ln y - \mu_{\ln S_a(T_1)}) / \sigma_{\ln S_a(T_1)}$ is the deviation (in log terms) of $S_a(T_1)$ away from its expected value (given M =m and R=r). The *conditional* median spectral ratio or shape is Eq. 6 divided by y or

$$\eta_{R_{2/1}} = \frac{\eta_{S_a(T_2)}}{\eta_{S_a(T_1)}} \exp[-\varepsilon_{S_a(T_1)}\sigma_{\ln S_a(T_1)}(1-\rho\frac{\sigma_{\ln S_a(T_2)}}{\sigma_{\ln S_a(T_1)}})] \tag{7}$$

Note that this conditional shape is the marginal median shape times an exponential that is negative and proportional to ε and (1-ρ) under the reasonable assumption that the two σ's are about equal. For the larger values of y of most engineering interest ε is positive and the conditional median shape at other periods is *below* the one we expected "on average". The degree to which it is below depends on how far the selected level y of $S_a(T_1)$ is above its median and how weakly correlated the two spectral ordinates are. This correlation decays with separation between the two periods[16]. The implication is that the conditional median spectral shapes of interest are somewhat peaked around T_1 relative to the median shape. This in turn means that if the response of the first mode is higher than expected the response of the second mode will be relatively less than one would expect otherwise. These conclusions have been verified empirically (Baker 2005c). Fig. 1 shows plots of predicted spectra, including one median spectrum and two that are conditional on Sa(0.8s) being at the 1.5 ε level. It also shows these three spectra scaled to a common value of Sa(0.8s). Note that the two positive ε spectra are more peaked than the median spectrum and have very similar shapes despite the half-unit magnitude difference in their causative earthquake.

If the structure has an important second-mode contribution to θ, a first-order requirement towards estimating well $G_{\theta|IM}(x|y)$ would be to select records that have approximately this median shape value. Note that for higher values of y that this shape is not the median shape. Therefore the practices of using the UHS shape or the median shape of the dominant M and R scenario[17] are not as accurate as they might be for the rare events of common engineering

[15] A recent study of this correlation coefficient is to be found in (Baker 2005b). It depends primarily on the ratio T_1/T_2 and is effectively independent of magnitude and distance.

[16] For separation of modal periods by a factor of 3 the value of ρ might be about 0.6 and $\sigma_{\ln Sa}$ about 0.7 implying for ε equal 1 the shape is about 25% lower and for ε equal 2 the shape is 40% lower. The effect on θ will depend on the importance of the second mode to the response quantity in question.

[17] For our special case here that the seismic threat is limited to single {M,R} scenario, the UHS shape and the median shape (given these M and R values) are virtually identical (differing only to the degree that the log standard deviations are different at different periods).

interest.

A direct way to select the records in this simple linear 2-DOF case is to calculate the spectral *ratio* at T_1 and T_2 for all the records and select records such that their median is about that given in Eq. 7. Then they may be scaled such that $S_a(T_1)$ equals level y and the dynamic analyses run to find $G_{\theta|IM}(x|y)$. Note in particular that for this linear 2-DOF case the records selected will need to change as y changes, that the M and R of the records are immaterial per se (only their spectral ratio matters), and that the selected records may be scaled to any degree without loss of accuracy.
Even in the linear 2-DOF case we have made simplifications. In the record selection we have tried only to match the conditional central value not the variability of the spectral ratio. This is also quite feasible but will not be pursued further here. Further if we do not have a single {M,R} pair the determination of this conditional median shape is not trivial. In principle there must be a weighting over all $\{m_i,r_i\}$ pairs[18]. It may be sufficiently accurate in some cases to use disaggregation to determine a single dominant {M,R}and then apply the reasoning above as if it were the only threat. This would already be a step beyond current practice.

A second more indirect way to select the records for this case is to chose records (from the general magnitude range; R is less critical as it has little mean effect on strong motion spectral shape) that have the required level of ε (consistent with the level y), and then scale them to the correct level of $S_a(T_1)$. This too should on average at least capture the more peaked spectral shape associated with the higher levels of y and ε of primary interest.

More General Structures

While the conclusions from the simple 1 and 2-DOF linear structure are in principle limited to these simple cases, they suggest several generalizations that can be made, which have been verified in recent studies with nonlinear SDOF and MDOF frame models (e.g., Iervolino 2005, Baker 2005c, Tothong 2005). The objectives in record selection and scaling for accurate estimation of $G_{\theta|IM}(x|y)$ are to capture primarily the proper general amplitude (via the IM level) and secondarily the spectral shape given that IM level. For the common IM $S_a(T_1)$ once the first objective is met the estimation of $G_{\theta|IM}(x|y)$ is fairly robust with respect to the records selected as long as the structure is first-mode dominated and only moderately nonlinear.

For other structures it can be important to select the records to capture the appropriate spectral shape. This was demonstrated for higher modes ($T_2/T_1 < 1$), but it is clear from our understanding of nonlinear dynamic behavior that it is equally important to reflect properly the longer periods when the structure experiences substantial “softening”. While there are no direct ways (similar to that used above for the 2-DOF case) to identify one or a few unique longer periods to focus upon and analyze, it should be clear that an objective of matching the conditional median spectral ratio would apply to each period of interest. Hence it follows that the entire *conditional* median spectral shape is logical first-order target for record selection for all structures. It should be re-emphasized that this shape is not the same as the median or UHS shape unless the level of y is near the median value of the IM $S_a(T_1)$, and that this shape will change as y changes, being most different from the median shape for large, rare values of the IM,

[18] As mentioned in footnote 14 such information will become available in time.

which is when strong degrees of nonlinearity may occur. It is the author's belief, while not proven here, that for more general structures (as was shown here for the 1 and 2-DOF linear structures) it is not critical to capture the "causative" magnitude[19] if the spectral shape itself has been selected well. Magnitude is primarily just a proxy for the median shape.

We saw above that arbitrary levels of scaling of the records was not a cause of response bias in the 1 and 2-DOF linear structures, provided (in the 2-DOF case) that the spectral shape was correct. It is the author's experience that, with this same proviso, this conclusion is more generally true (e.g., Shome 1998, Baker 2005c).

It should be noted that the schemes discussed above are based on the common use of linear first-mode period spectral acceleration as the IM. It has been found that other choices of the IM may provide even more robustness with respect to record selection (Luco 2002, Luco 2005, and Tothong 2005), just as $S_a(T_1)$ provides more such robustness than PGA. These new IM's are based on inelastic spectral acceleration. Another major objective of seeking improved IM's is the reduction of the number of records and analyses needed to achieve a specified level of confidence in the estimate (recall it was set here as a standard error of estimation of $\lambda_\theta(x)$ of about +/- 30%, or of the conditional mean of θ of about +/- 10%). This is achieved by finding "better response predictors", i.e., IM's that reduce the variance of θ given IM = y for various levels of y (i.e., degrees of structural nonlinearity). This so-called IM efficiency (Luco 2005) issue has not been addressed here.

Further this study of the record selection problem presumes that the model of the structure is available at least to the level of knowing the general range of its first-mode period. (Because of the high correlation between two comparatively nearby periods there is little loss of accuracy or efficiency if the period T of the Sa(T) used as the IM is some distance from that of the final first-natural period of the structural model. (Of course whether that model estimates well the first-natural period of the real structure is another problem, which does not influence how we should best analyze the model we have.) While the general principles above hold for all cases the judgments stated as to accuracy and effectiveness depend on the author's experience with building-like structures, which at least in the linear range tend to have no more than two or three most-important elastic modes. Other cases, including those in which the record selection is done (for good or bad reasons) without knowledge of the structure or those where there are many very important response measures sensitive individually to different portions of the input spectrum, have not as yet been studied in this way. Short of using different IM's, records, and analyses for the different subsets of these cases (a strategy which is permitted in the nuclear arena, e.g, NRC 1997), it is clear that some compromise well have to be made with respect both to efficiency (implying larger sample sizes or larger standard errors) and perhaps to the accuracy (or record selection robustness).

---

[19] Exceptions may be when the response is duration sensitive and magnitude serves as a proxy for duration. The peak displacements of nonlinear framed structures do not seem to be duration sensitive even when strength degradation is involved.

Conclusions

We conclude that the selection of records for use in nonlinear seismic time history analysis of MDOF models of structures can benefit from starting from a defined structural objective, here estimation of the mean annual frequency of some structural response measure, θ, exceeding level x, i.e., $\lambda_\theta(x)$, and then asking how that might be directly and/or indirectly estimated in various ideal and simplified cases. Discussion of the ideal case in which 10,000+ years of recordings have been made at the site reveals that numerous (order 1000) records and time history analyses will be required (save, perhaps, special techniques to reduce this number). Recognizing that the recorded accelerograms must come from catalogs of data recorded at many sites and caused by many sources, it becomes clear that record selection must allow in some way for the failure of such data to reflect the relative frequency with which various events at various distances will affect the site. In short in this form the record selection problem is very site-seismicity dependent. This complicates the selection problem and increases the number of analyses necessary for accurate estimation of $\lambda_\theta(x)$.

Turning to an intensity-measure-based formulation of the analysis and estimation problem, one finds that the PSHA analysis leading to the IM hazard curve, $\lambda_{IM}(y)$, captures much of the site-specific seismicity issue and of the variability in θ. Now one needs to select records to estimate $G_{\theta|IM}(x|y)$, the CCDF of the θ conditional on IM = y. Even though this CCDF may have to be evaluated at several levels of y, this approach leads to significantly reduced record selection and sample size needs. But again the ideal problem is compromised by the need to select records from the available record set rather than a site-specific catalogue. Two, simplest structural dynamic models, namely linear 1- and 2-DOF systems, are discussed with the objective of understanding the benefits and challenges of the record selection and scaling problem for "real" (nonlinear MDOF) structural models. The conclusions are that, while the sample sizes may be smaller with the IM-based procedure, to the degree that the structure is not first-mode dominated or not nearly linear, care may be needed in selection of the records (and particularly with respect to their spectral shape), to insure that a bias is not introduced in the estimation of $G_{\theta|IM}(x|y)$ and hence in $\lambda_\theta(x)$. In particular it appears that the spectral shape needs to reflect properly the level of ε, which is a measure of the $S_a(T_1)$ IM level *relative to* its expected value (given the M and R scenarios of primary interest at the site). This shape is, for the positive ε's of engineering interest, more peaked than the UHS or median spectral shape (given a dominant scenario {M,R} pair). *Provided this shape is captured* it appears that the M and R selection criteria may be significantly relaxed and that scaling records to match the level y of the IM will not induce significant bias. These conclusions are supported by current research on nonlinear MDOF frames (e.g., Baker 2005a, Baker 2005c, Iervolino 2005).

It is noted that, while the focus here has been on estimating $\lambda_\theta(x)$, the conclusions here apply to "more deterministic" current-code-practice objectives, such as estimating the mean of θ "given the 2% in 50 ground motion". Since this concept in quotes exists only for a scalar ground motion measure (and not, for example, for an entire spectrum), it can be taken here to mean the 2% in 50 $S_a(T_1)$ IM level. Setting y equal to that level y* for which $\lambda_{IM}(y^*) = 0.0004$, this problem can be stated as estimating the mean of the distribution $G_{\theta|IM}(x|y)$, i. e., the conditional mean of θ given $S_a(T_1) = y^*$. It will be recalled that concerns about the required sample sizes and about biasing that distribution were cast above in terms of that mean. More modern codes are based on a target value of $\lambda_\theta(x)$ (rather than $\lambda_{IM}(y)$). Therefore the discussion above applies directly to them.

Acknowledgements

This simple paper depends strongly on the far more elegant and complete PhD thesis work of Dr. Jack Baker (Baker 2005c). I want to acknowledge his primary contributions, and the many good discussions with him, Polsak Tothong, and Iunio Iervolino on this subject. This work was supported 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). 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.

References

American Society of Civil Engineers (2005). *Seismic design criteria for structures, systems, and components in nuclear facilities*, Structural Engineering Institute, Working Group for Seismic Design Criteria for Nuclear Facilities, *ASCE/SEI 43-05*, Reston, Va., 81 pp.

Anderson, J. and J. N. Brune (1999). "Probabilistic seismic hazard analysis without the ergodic assumption", *Seismological Research Letters*, 10 (1) 19-28.

Baker, J. W. and C. A. Cornell (2005a). "A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon", *Earthquake Engineering & Structural Dynamics*, 34, 1193-1217.

Baker, J. W. and C. A. Cornell (2005b) "Correlation of response spectral values for multi-component ground motions", Accepted for publication, *Bulletin of the Seismological Society of America.*

Baker J.W. and C. A. Cornell (2005c). Vector-valued ground motion intensity measures for probabilistic seismic demand analysis, *John A. Blume Earthquake Engineering Center Report No. 150*, Stanford University, Stanford, CA.. http://blume.stanford.edu/Blume/Publications.htm.

Bazzurro, P., C. A. Cornell, N. Shome, and J. E. Carballo (1998). "Three Proposals for Characterizing MDOF Nonlinear Seismic Response," *Journal of Structural Engineering, ASCE* 124 (11), 1281-1289.

Bazzurro, P., and C. A. Cornell (1999). "On Disaggregation of Seismic Hazard," *Bulletin of Seismological Society of America,* 89 (2), 501-520.

Bazzurro, P. and C. A. Cornell (2002). "Vector-Valued Probabilistic Seismic Hazard Analysis (VPSHA)," *Proceedings 7th U.S. National Conference on Earthquake Engineering*, Boston, MA.

Benjamin J.R. and C. A. Cornell (1970). *Probability, Statistics, and Decision for Civil Engineers*, McGraw-Hill, New York, 684 pp.

Han S.W. and Y. K. Wen (1997). Method of reliability-based seismic design: Equivalent nonlinear systems, *Journal of Structural Engineering* 123 (3), 256-263.

Iervolino, I., and C. A. Cornell (2005). "Record selection for nonlinear seismic analysis of structures", *Earthquake Spectra* 21 (3), 685-713.

Jalayer F., (2003). Direct probabilistic seismic analysis: Implementing non-linear dynamic assessments, *Ph.D. Thesis*, Stanford University, Stanford, CA, 244 pp. http://www.stanford.edu/group/rms/

Jalayer F. and C. A. Cornell (2003). A technical framework for probability-based demand and capacity factor design (DCFD) seismic formats, *Pacific Earthquake Engineering Research Center Report PEER 2003-08*, University of California at Berkeley, Berkeley, CA, 106 pp.

Jalayer F., J. L. Beck, and K. A. Porter (2004). Effects of ground motion uncertainty on predicting the response of an existing RC frame structure, *Proceedings 13th World Conference on Earthquake Engineering*, Vancouver, Canada.

Luco, N. and C. A.Cornell (2005). "Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions," *Earthquake Spectra*, Submitted.

Luco, N., M. Mai, C.A. Cornell, and G. Beroza (2002). "Probabilistic seismic demand analysis at a near-fault site using ground motion simulations based on a kinematic source model," *Proceedings 7th U.S. National Conference on Earthquake Engineering*, Boston, MA.

Nuclear Regulatory Commision (1997). "Identification and characterization of seismic sources and determination of safe shutdown earthquake ground motion", *Regulatory Guide 1.165*. www.nrc.gov/reading-rm/doc-collections/reg-guides/power-reactors/active/01-165/ (accessed 8/17/2005)

PEER strong motion database (2005). http://peer.berkeley.edu/smcat/ (accessed 6/29/2005).

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

Somerville P.G. and H. K. Thio (2005). Probabilistic vector-valued seismic hazard analysis for near-fault ground motions, *Southern California Earthquake Center Project Report* (in preparation).

Stewart J.P., S. J. Chiou, J. D. Bray, R. W. Graves, P. G. Somerville, and N. A. Abrahamson (2002). Ground motion evaluation procedures for performance-based design, *Soil Dynamics and Earthquake Engineering*, 22 (9-12), 765-772.

Tothong P. and C. A. Cornell (2005). Near-fault ground motions for seismic demand analysis (manuscript in preparation).

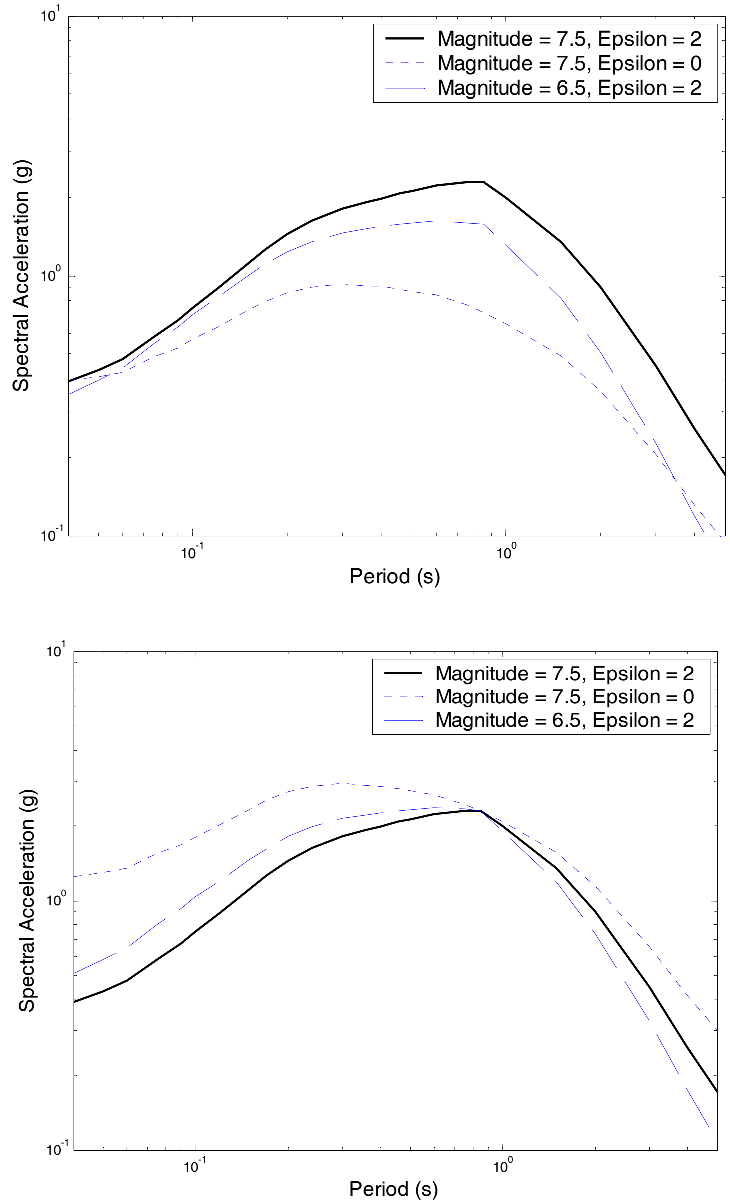

Figure 1. (Upper) expected response spectra for three scenario events; (lower) expected response spectra for three scenario events, scaled to have the same $S_a(0.8s)$ value.

Source: Baker 2005a.