Simulation records (2) are obtained via modeling of the seismological source and may account for path and site effects. These methods range from stochastic simulation [Boore, 2003] of point or finite sources to dynamic models of rupture.

Synthetics may be the only way to obtain appropriate records for rare scenarios, such as large magnitude events "close" to the site and give the benefit that one can produce from them large samples of nominally similar events. However, effort should be employed to insure that their spectra are appropriate for nonlinear analysis, e.g., non smooth [Cornell, 2004]. Moreover, they often require setting of some rupture parameters, such as the *rise-time*, which are hard to determine. *S*ome state-of-the-art simulation methods seem to overcome these shortcomings, but they are not yet readily available to engineers.

Finally, of type (3) are ground-motion records from real events. The availability of on-line, user-friendly, databases of strong-motion recordings, and the rapid development of digital seismic networks worldwide, have increased the accessibility to recorded accelerograms. However, due to the large variability in records representing a scenario, a number of points arise regarding the criteria for appropriate selection and manipulation of such records. In particular, an issue regarding the use of real recordings, whose spectra are generally non smoothed, is the selection of a set compatible with a code-specified spectrum. To overcome this, various approaches have been developed to manipulate real records to match a target spectral shape, either by frequency-domain or by time-domain modification methods such as the *wavelet* transform. The wavelet transform basically consists of using modulating functions, selectively located in time to modify the spectrum of the signal, where and when it is needed in order to match the target spectrum (see Hancock et al., 2006, for details). Although these methods produce records perfectly compatible with code's prescriptions and have the additional advantage of reducing the dispersion in the response, and hence the required sample size, some studies show that they may lead to a non conservative estimation of the seismic response [Carballo and Cornell, 2000; Bazzurro and Luco, 2003]. Therefore, earthquake engineering research has focused lately on the selection of real ground-motions for non-linear structural analysis and relatively simple and effective procedures have been developed to link records to the hazard at the site.

In seismic codes, the guidelines about preparation of ground-motion input for dynamic analysis are generally poor, as also pointed out by Bommer and Ruggeri [2002]. This is partially because research on the topic is developing fast and at least a few years are required by regulations to take it in. For example, the code-based prescriptions for records often require compatibility with a smooth design acceleration spectrum together with few other minor requirements. Eurocode 8 (EC8) [CEN, 2003], in particular, allows employment of all three kinds of accelerograms listed above as an input for seismic structural analysis. The EC8 prescriptions ask for matching of the average spectral ordinates of the chosen record set to the target code-based spectral shape. The set has to consist of at least seven recordings (each of which includes both horizontal components of a recorded motion if spatial analysis is concerned) to consider the mean of the response. Otherwise, if the size of the set is from three to six, the maximum response to the records within the sets needs to be considered. Little, if any, prescriptions are given about other features of the signal. Therefore, the code requirements seem to have been developed having spectrum-compatible records in mind. On the other hand, real accelerograms are becoming the most attractive option to get unbiased estimations of the seismic demand.

The study presented herein investigates the feasibility of finding real record sets complying as much as possible with EC8 spectra. EC8 does not provide anchor values (a_g) for its non dimensional spectral shapes, leaving it up to the European national authorities to determine the values of a_g, which is associated to the Peak Ground Acceleration (PGA) on rock with a certain probability of exceedance at the site of interest. The a_g values employed herein correspond to the Italian case, where the seismic territory is divided into four zones representing different hazard levels, where seismic resistant design is mandatory only in the upper three zones.[1]

The chosen ground-motion spectra dataset is extracted from the European Strong-Motion Database, which contains accelerograms from both European and Mediterranean events. Original spectra from this database have been combined in all possible suites of seven in order to find EC8 compliant sets of un-scaled records. Moreover, sets of scaled code-compatible accelerograms were also considered in order to reduce to record-to-record variability in the response, and to obtain sets which are independent on the anchoring value of the code spectrum.

Finally, sets compatible with Eurocode 8 spectra, for plane and spatial analysis of buildings, are found and some of them are discussed herein. The selected records refer to rock or stiff soil site classes and are available on the internet on the website of the Italian consortium of earthquake engineering laboratories: *Rete dei Laboratori Universitari di Ingegneria Sismica – ReLUIS* [http://www.reluis.it/]. On the same website, similar results and discussion for the selection of ground-motions suitable for dynamic analysis and compatible with the recent Italian seismic code prescriptions (slightly different from those of EC8), are also given [Iervolino *et al.*, 2006a].

2. Current Best Practice and Critical Issues in Record Selection and Manipulation

Among the possible approaches, reviewed by Beyer and Bommer [2007], in selecting real accelerograms for assessing the nonlinear demand of structures, the current state of best practice [Cornell, 2005] is based on first disaggregating the seismic hazard at the site [Bazzurro and Cornell, 1999], by causative magnitude (M) and distance (R), for the level of spectral acceleration (at the first mode period of the structure) at a specified probability (say a 10% 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 and site conditions, 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, duration, instrument housing, etc. Finally, they are scaled to match, in some average sense, 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.[2] Based on the studies that have investigated this procedure, there is some evidence that all this care taken about the selected records' earthquake properties may be not justified [Iervolino and Cornell, 2005]. That is, it is not proven that record characteristics such as M and R significantly influence linear or nonlinear response conditioned to first mode spectral acceleration or another *sufficient*[3] ground-motion intensity measure (IM). Moreover, the scaling of records to match some spectral value does not seem to bias the response estimate if the deviation from the median ground-motion prediction relationship effect is accounted for appropriately [Baker and Cornell, 2006a].

[1]The a_g values for the Zones 3, 2, and 1 are 0.15 g, 0.25 g, and 0.35 g, respectively. These values are related to the probabilistic seismic hazard analysis (PSHA) [McGuire, 1995] for the site of interest. In fact, if the PGA (on rock) with a 10% exceeding probability in 50 years falls in one of the intervals]0.25g, 0.35g],]0.15g, 0.25g], or]0.05g, 0.15g], then the site is classified as Zone 1, 2, or 3, respectively [OPCM 3519, 2006].

[2]Many authors have recently questioned the use of UHS as target spectrum, see for example Baker and Cornell [2006a].

[3]Sufficiency of an intensity measure is discussed in detail in the next section.

When following Eurocode 8 criteria (described next), these procedures for record selection are not readily applicable because: (i) the code spectrum is related to the hazard for the site of interest only through the anchoring value, which is related to the PGA with a 10% exceedance probability in 50 years on a rock site, therefore it is not possible to apply common disaggregation procedures or to match any source parameter if a site-specific probabilistic seismic hazard analysis (PSHA) is not available; (ii) the requirement to match, in the average, the code spectrum in a broad range of periods seems to be very hard to satisfy. In the following, a brief review of recent developments and relevant literature on the topic of record selection for dynamic analysis is given because it may matter to understand the applicability of the results found and will help to discuss the EC8 prescriptions.

2.1. Sufficiency and Efficiency of a Ground-Motion Intensity Measure

A *sufficient* IM renders the structural response, conditioned on that IM, independent, of earthquake ground-motion characteristics such as magnitude and distance [Cornell, 2004]. At the same time, a certain IM is defined as *efficient* if the structural response, conditioned on IM, has comparatively small dispersion. The spectral acceleration (Sa), at the fundamental period of oscillation of the structure, is often implicitly assumed to be both a sufficient and efficient IM. This is in part due to the availability of Sa hazard curves; however, for inter-story drift response, Sa is at the very least more sufficient and efficient than PGA.

First-mode spectral acceleration has also been proven to be sufficient in respect to duration, at least for single degree of freedom (SDOF) structures [Iervolino *et al*., 2006b]. In fact, although there is a debate on the influence of duration in seismic assessment of structures, as reviewed by Hancock and Bommer [2006], duration has been found to be statistically insignificant to displacement ductility demand; conversely, it strongly affects, as expected, other demand parameters accounting for cyclic behavior such as hysteretic ductility or equivalent number of cycles. Therefore, at least for the purposes of displacement-related demand assessment, it seems that one should not take too much care in selecting records from a particular duration bin given that they have (or are scaled to) a common Sa level (e.g., matching of the target spectral shape at some frequency).

A sufficient and efficient IM also allows the estimation of the response (i.e., median) requiring a smaller sample size to get a given standard error. It has been demonstrated, in fact, that if Sa is concerned, the uncertainty on the estimation can be dramatically reduced if records are scaled to a common Sa level [Shome *et al*., 1998]. However, it is worth noting that it has also been recently demonstrated that Sa may not be particularly efficient, nor sufficient, for some structures. If long periods of oscillation are called into question, the higher modes typically play a larger role in the seismic response and Sa has less prediction power than for first-mode dominated structures. It may indeed be insufficient because it is not able to capture the spectral shape in a range of frequencies where the latter depends on the magnitude. For soft-soil or near-source records Sa may also be insufficient. At the same time, PGA may be a better IM for peak floor acceleration, which is an important response variable for non structural response, since non structural elements are often sensitive to applied inertia forces. The lack of efficiency, or insufficiency, of first mode Sa may be explained by the iconoclastic statement: “only spectral shape matters” in the estimation of nonlinear seismic response of structures. This means that in the case of systematic spectral shape deviations, Sa may be found to be insufficient. Several studies propose alternative IMs trying to capture spectral shape in a range of interest to some structural types. They include scalar and vectors, linear and nonlinear quantities; however, they are still not included in common practice.

2.2. *Epsilon*

It has been briefly reviewed above why in seeking for characteristics to mirror in the record selection one should look to any systematic effect on spectral shape. For example, it is prudent to avoid selecting records from soft soil sites or from near-source records showing directivity effects. Baker and Cornell [2005, 2006a] recently demonstrated that one source of systematic effect is that of the deviation of a record's Sa from the value predicted by the ground-motion prediction equation. That deviation is called *epsilon*[4] or "normalized residual." High epsilon values are associated with peaks in the spectrum (Fig. 1), and hence with more benign nonlinear structural behavior. In fact, during the shaking the effective period of the structure lengthens descending the peak toward a less energetic portion of the frequency content.

Even though some researchers believe that epsilon is not an intrinsic ground-motion feature, PSHA disaggregation for epsilon often shows that high IM levels, contributing directly to rare maximum interstory drift ratio (MIDR) levels, are associated with high values of epsilon. Therefore, when selecting records for analyses at these high IM levels, one should consider choosing them among those having the *right* epsilon, in order to have the *correctly deviating* spectral shape around the period of interest, for a more efficient

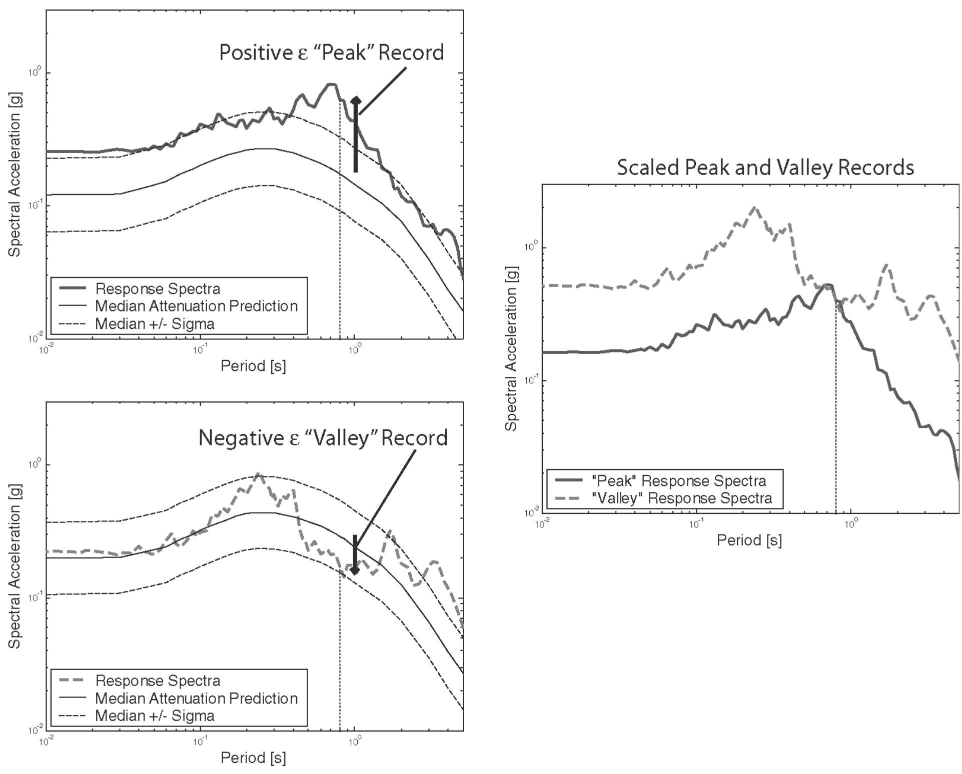

FIGURE 1 Scaling a negative ε record and a positive ε record to the same spectral acceleration at the period of 0.8 s. Courtesy of Jack W. Baker; see Baker and Cornell [2005] for details.

[4] Epsilon (ε) is defined as the difference between the log of the spectral acceleration, at a given period, of a record and that predicted by an ordinary ground-motion prediction equation divided by the standard deviation of the residuals.

and unbiased estimation of structural response. This is more important than matching records with scenario M and R values. It is also worth noting that the epsilon issue affects the scaling procedure; in fact, for example, scaling down a positive epsilon record would introduce an un-conservative bias in the demand estimation because, due to the lengthening of the period during the shaking, the structure will be sensitive to a part of the spectrum which is away from the peak; conversely, scaling up a negative epsilon record could lead to an overestimation of the seismic response (Fig. 1).

A method has also been proposed [Baker and Cornell, 2006a] to develop a target spectrum which accounts for the effect of magnitude, distance, and epsilon. This spectrum allows the selection of records that only have a spectral shape that matches the mean spectrum from the causal event, without taking care of appropriate magnitude, distance, and specific epsilon. The proposed target spectrum is compared to an UHS, and seen to be more appropriate for obtaining unbiased estimates of structural response.

2.3. Consistent Sa

In performing seismic assessment of structures via dynamic analysis it is important to bear in mind that structural engineers and seismologists sometimes intend Sa differently. This mismatch is due to the decomposition of ground-motion by projection along two directions [Baker and Cornell, 2006b]. For the aims of nonlinear seismic assessment of structures, Sa is considered as the one along a single axis. Conversely, seismologists may compute ground-motion prediction equations using the geometric mean of the spectral accelerations in the two directions; using one arbitrary component would lead to a larger dispersion of hazard curves. Both uses of Sa are legitimate, but inconsistent if combined for the probabilistic seismic assessment of structures. Therefore, it is preferable to define the same Sa in both the hazard and response. This means either that in the seismic risk analysis of structures one should use hazard curves that use one-component Sa, or estimating structural response using the geometric mean of the two components as an IM. This latter method has the advantage of not requiring new ground-motion prediction equations for hazard analysis. However, it will introduce additional dispersion into the response prediction and Sa will result less efficient. Alternatively, if the structural response is estimated using a single axis Sa, while hazard refers to the mean of the two components, the dispersion of the response may be inflated, as proposed by the cited authors, to reflect that which would have been seen if the mean Sa had been used as the intensity measure.

2.4. Near-Source

Finally, it should briefly be mentioned that a site located close to the source of a seismic event may be in a geometrical configuration, in respect to the propagating rupture, which may favor the constructive interference of waves (synchronism of phases causing building up of energy) traveling to it, which may result in a large velocity pulse. This situation, for dip-slip faults, requires the rupture going toward the site and the alignment of the latter with the dip of the fault, whereas for strike-slip faults the site must be aligned with the strike; if these conditions are met the ground-motion at the site may show *forward directivity* effects [Somerville *et al*., 1997]. Parameters driving the amplitude of the pulses are related to the above-discussed rupture-to-site geometry, while empirical models positively correlating the earthquake's magnitude to the period of the pulse have been proposed by seismologists [Somerville, 2003]. Pulse-type records are of interest for structural engineers because they: (1) may induce unexpected demand into structures having the fundamental period equal to a certain fraction of the pulse period; and (2) such demand may not be adequately captured by the current, best-practice, ground-motion intensity measures such as first mode spectral acceleration.

Common record selection practice and classical PSHA do not apply in the near-source. In fact, the latter requires ground-motion prediction relationships able to capture the peculiar spectral shape driven by the pulses, while the former should produce record sets reflecting the pulse features compatible with the near-source PSHA. Extended discussion and results on the topics of near-source hazard analysis and seismic assessment in near-source conditions may be found in the work by Tothong [2007].

3. Eurocode 8 Prescriptions for Record Selection

Eurocode 8, part 1, outlines the requirements for the seismic input for dynamic analysis in *Sec. 3.2.3:*[5] *The seismic motion may be represented in terms of ground acceleration time-histories and depending on the nature of the application and on the information actually available, the description of the seismic motion may be made by using artificial accelerograms (see 3.2.3.1.2) and recorded or simulated accelerograms (see 3.2.3.1.3).*

The set of accelerograms, regardless if they are natural, artificial, or simulated, should match the following criteria:

a. *a minimum of 3 accelerograms should be used*;
b. *the mean of the zero period spectral response acceleration values (calculated from the individual time histories) should not be smaller than the value of a_g S for the site in question;*[6]
c. *in the range of periods between 0,2T_1 and 2T_1, where T_1 is the fundamental period of the structure in the direction where the accelerogram will be applied; no value of the mean 5% damping elastic spectrum, calculated from all time histories, should be less than 90% of the corresponding value of the 5% damping elastic response spectrum*.[7]

Some duration prescriptions are given for artificial accelerograms, while recorded or simulated records should be *adequately qualified with regard to the seismogenetic features of the sources and to the soil conditions appropriate to the site, and their values are scaled to the value of a_g S* (PGA) *for the zone under consideration*. Regarding the former part of the sentence: it has to be noted that the code spectrum is not related to any specific feature of the source and, therefore, this prescription could not be accounted for herein. The latter part was not considered as well because: (1) it is not very clear what scaling the values of the records means; (2) if it means to scale the PGA of the individual records to the PGA value of the code, then the condition (b) above seems to make this statement useless (in fact, many codes such as the Italian one do not have this statement, although they have very similar prescriptions including (b) above).

According to the code, in the case of spatial structures, the seismic motion shall consist of three simultaneously acting accelerograms representing the three spatial components of the shaking, then *3* of condition (a) shall be considered as the number of *groups* of

[5]In the rest of the article, all calls and verbatim citations of Eurocode 8 will be simply indicated in italic.

[6]Many national codes in Europe have the EC8 as a main reference. The recent Italian seismic code [OPCM 3274, 2003], for example, has very similar prescription for record selection. However, this (b) criterion is not present.

[7]The upper limit accounts for the lengthening of period due to the nonlinear structural behavior, while the lower considers the contribution of higher modes to structural response. The recent Italian seismic code prescribes the lower period range limit as 0.15 s.