Strong ground motion record selection for the reliable prediction of the seismic collapse capacity of structures[*]

M. Mousavi & M. Ghafory-Ashtiany
International Institute of Earthquake Engineering and Seismology, Iran

A. Azarbakht
Arak University, Iran

ABSTRACT:
How to select a limited number of strong ground motion records (SGMR's) is an important challenge for the seismic collapse capacity assessment of structures. The goal of this paper is to select, from a general set of SGMR's, a small number of subsets such that each can be used for the reliable prediction of the seismic collapse capacity of a particular group of structures, i.e. of SDOF systems with a typical behaviour range. The multivariate statistical analysis (MSA) and the principle component analysis (PCA) are two key aspects of the proposed methodology. The methodology has been validated by analysing a three-storey-reinforced concrete structure by means of the proposed subsets, as well as the general set of SGMR's. Also, the proposed subset shows good agreement with the general set for the prediction of the IM in the full range of EDP, i.e. maximum drift.

Keywords: strong ground motion record (SGMR), seismic collapse capacity, structure group

1. INTRODUCTION

One of the challenges in the non-linear response-history analysis of sophisticated structural models is how to select a limited number of SGMR's. The selection of appropriate SGMR's needs to be performed with the goal of accurately estimating the response of a structure to a specified ground motion intensity, as measured by the spectral acceleration corresponding to the first mode period of the structure, Sa(T1). The current code-based method of record selection (e.g. [1]) is based on a consideration of the magnitude and distance of the SGMR's, while matching the mean response spectrum to the uniform hazard spectrum (UHS) as a target spectrum. Since no single earthquake is likely to produce a spectrum as high as the UHS spectrum, the code-based procedure for record selection is usually conservatively biased [2-4]. Reduction of this bias and the variance of the resulting structural response can be achieved by considering the spectral shape in the record selection [4]. It has been demonstrated that spectral shape can be indicated by epsilon, which is defined as a measure of the difference between the spectral acceleration of a record and the mean obtained from a ground motion prediction equation at a given period [5]. It can therefore be concluded that one method to account for the spectral shape effect is through the selection of a set of SGMR's that is specific to the structure's fundamental period and the site hazard characteristics [6]. This selection presents a significant challenge when assessing the seismic collapse capacity of a large number of structures or when developing a systematic procedure, since it implies the need to assemble specific ground motion sets for each structure. An alternative method has been proposed in [6], whereby a general set of SGMR's is used to simulate collapse, and the resulting collapse capacity is adjusted in order to take into account the spectral shape effects that are not reflected in the selection of the general set. The

---

[*] This paper is based on a manuscript which is submitted to the Earthquake Engineering and Structural Dynamic journal, 2010.

major difficulty of this method is that it implies the need to apply a relatively large number of ground motion records for the collapse assessment of the structures involved.

The main object of this paper is to find a proper solution for a reduction in the number of SGMR's needed for a non-linear response-history analysis. The proposed approach is intended to suggest a few subsets from a general set of SGMR's, and to use each subset for the collapse simulation of a specific structure group with a typical range of behaviour. The main criterion for selection of each subset is the similarity of the subset to the general set for the prediction of the collapse capacity of the structures belonging to the relevant structure group. It is important to note that suggestion of subsets is done just once, and that the selected subsets can be used a priori for the prediction of the collapse capacity of any arbitrarily selected first-mode dominated structure.

2. METHODOLOGY

The general SGMR set can be chosen from a catalogue, based on the seismological aspects of a considered earthquake scenario [1]. The second set of SGMR's, which is hereafter called a subset, can be selected from the general SGMR set, but has a smaller number of SGMR's. The similarity of the structural seismic collapse capacity response, based on the two different SGMR sets, can be measured by comparing the corresponding probability density functions as described in the following sub-section.

2.1. Quantification of the similarity of two sets of SGMR's

Assume that $X_1, X_2, ..., X_n$ and $Y_1, Y_2, ..., Y_m$ are samples from a population which are called, respectively, the general set and the subset. Based on Equation (1), two plausible ranges can be defined for the population mean by considering a 99% confidence interval ($\alpha = 0.01$).The degree of similarity of the two mean values can be quantified by defining two similarity indices. The first index, S_1, is the probability that the population mean determined from the general set falls into the plausible range of the mean which is obtained from the subset. The second similarity index, S_2, is defined as the probability that the population mean determined from the subset falls into the plausible range of the mean obtained from the general set. The similarity indices are computed, as shown in Figure 1, by integrating the related probability density functions over the associated ranges.

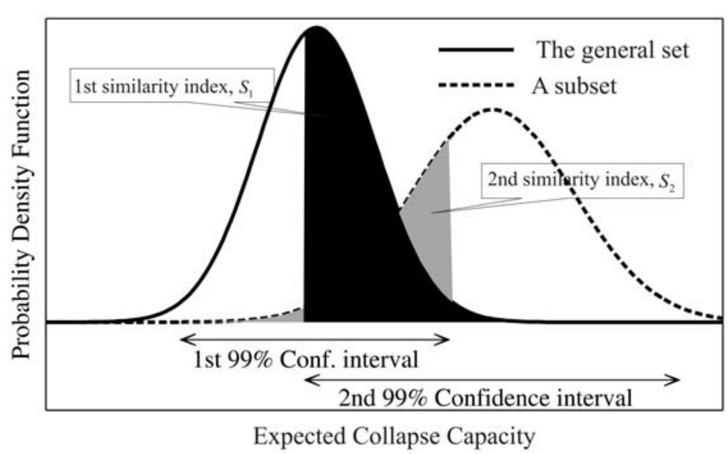

Figure 1. The definition of the two similarity indices for the univariate case.

The above-described approach for the assessment of the similarity of two sets of SGMR's can be extended to apply to an analysis of a set of SDOF systems categorized as a structure group. For this purpose, the univariate case, as shown in Figure 1, can be developed into the multivariate case by using the multivariate statistical analysis technique. As an example, the different confidence regions are shown schematically, for a given $p = 2$ dimensional sample, in Figure 2.

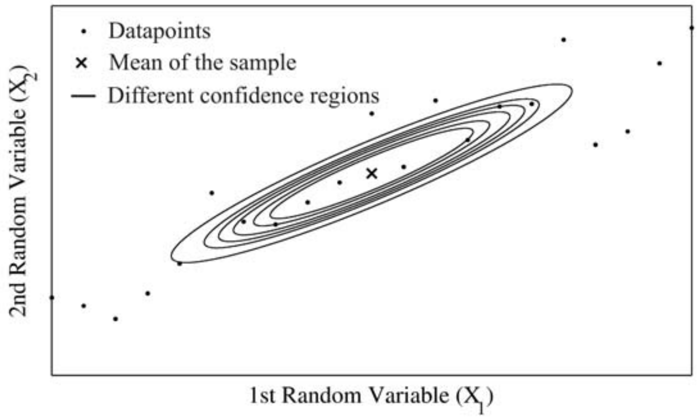

Figure 2. The different confidence regions for a two dimensional statistical sample.

The orientation of the confidence ellipses is proof of the correlation between the variables, i.e. between the structural seismic collapse capacities. So it is a reasonable inference that the p correlated variables must be analyzed jointly, which is a key aspect of multivariate statistical analysis [7]. The similarity of the two sets of SGMR's for a structure group of size p can be quantified by extending the previously-defined similarity indices to the multivariate case.

2.2. Selection of a near-optimal SGMR subset

The near-optimal subset of SGMR's can be selected from the general SGMR set by maximizing the similarity indices. Here, GA, as an efficient optimization search algorithm, is applied [8].

Generally, by increasing the size of the subset, the similarity indices will increase. On the other hand, by limiting the minimum value of the similarity indices to 90%, the minimum size of the subset can be calculated. For example, for the dataset used in Figure 3, the minimum number of SGMR's in the subset that guarantees similarity indices of at least 90% is six. The selected subset and the corresponding confidence region are shown in Figure 3.

The size of the subset will increase significantly when the size of structure group is increased. In the following section, a statistical solution is proposed in order to solve this problem.

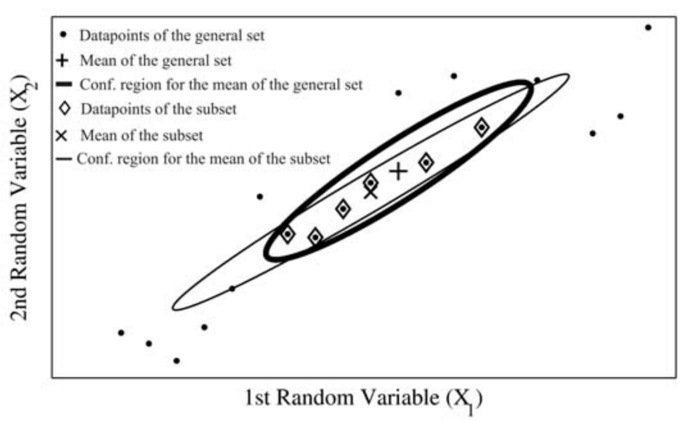

Figure 3. The mean values and the corresponding confidence intervals for the general and the subset of SGMR for a two dimensional problem.

2.3. Reduction in the size of the SGMR subset by PC (Principal Components) analysis

The role of PC analysis in the reduction of the subset size is shown in Figure 4. A high correlation between the two variables, as seen in Figure 4a, makes it possible to reduce the dimensions of the data-set. The transformed data is shown in Figure 4b, which confirms the independence of the PC's. The first PC, as seen in Figure 4c, contains 95% of the total variation in the data, so that neglecting the second PC is a rational decision. In this case, only three data-points are needed to supply the minimum 90% of the similarity indices (compared to the six SGMR's, as presented in section 2.2). The similarity indices, as well as the probability density functions, are presented, for the general set and for the subset of SGMR's, in Figure 4d. The good agreement between the probability density functions, associated with the general set and the near-optimal subset, confirms that the use of the first PC results in a reduction in the minimum size of the subset from six to three.

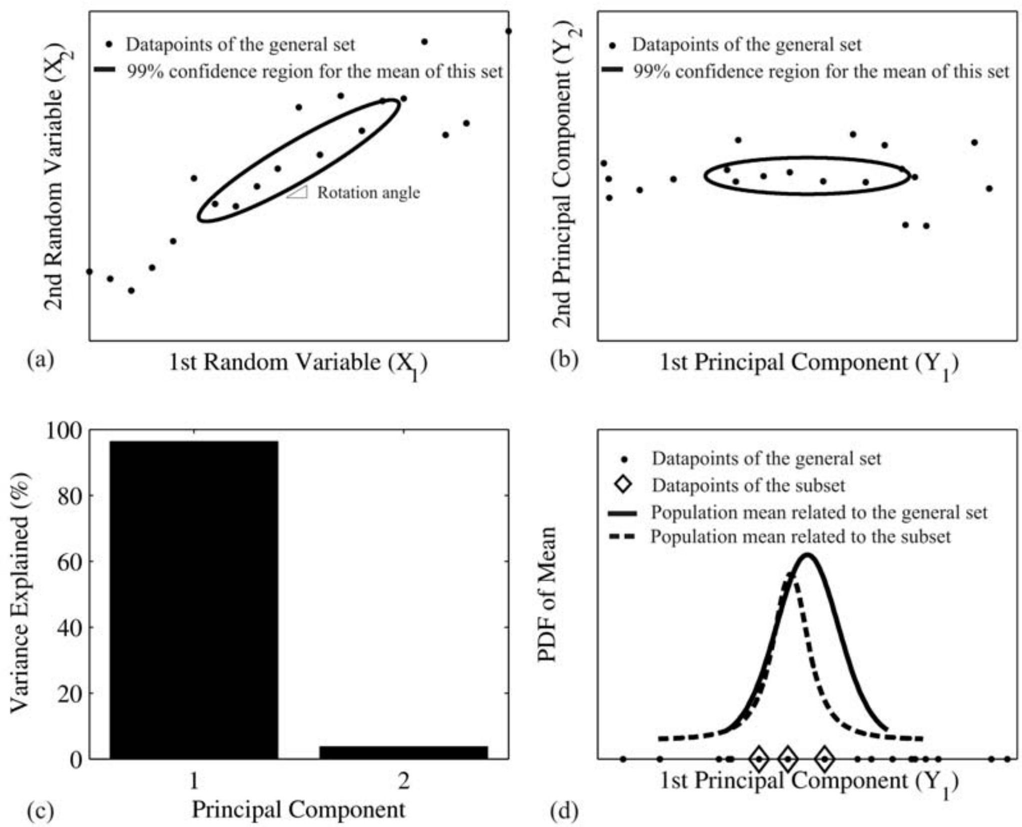

Figure 4. The Role of PC analysis in reduction of subset size: (a) a high correlation between the two variables can be observed, (b) the transformed data confirm the independency of the PC's, (c) the first PC contains 95% of the required information, and (d) there is good agreement between the probability density functions corresponding to the general and the subset.

3. SELECTION OF SGMR SUBSETS FROM THE GENERAL SET USED IN THE ATC63 PROJECT

The proposed methodology was applied to a specific far-field set consisting of 44 SGMR's. This general set was used in the Applied Technology Council Project (FEMA 2008) as a procedure to validate the provisions for seismic structural design [9]. A seismic collapse capacity database was established for 84 SDOF systems with periods ranging between 0.1 and 2.0 sec (T = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.25, 1.5, 1.75 and 2.0), six ductility values (μ_f = 2, 4, 6, 8, 10, and 12), and a mass proportional critical damping ratio equal to five percent.

As illustrated in Table 3, based on the results of these analyses, all of the 84 SDOF systems were grouped into six discrete groups, and eight near-optimal records were found for each group. As an interesting result, the optimum pattern of classification is just based on period and consequently the selected records are ductility independent. Table 2 shows the characteristics of the SGMR's which are proposed for each of the groups described in Table 1. In the next section, the proposed SGMR subset will be applied to estimate the seismic collapse capacity of a test structure, in order to investigate the efficiency of the proposed methodology.

Table 1. The near-optimal SGMRs for different period ranges.

Ground Motion Subset	Period	Variance Explained by the 1st PC	Variance Explained by the 2nd PC	SGMs ID	1st Similarity Index	2nd Similarity Index
I	0.1 ~ 0.3	59.6%	26.1%	3 – 8 – 14 – 20 – 21 – 24 – 27 – 28	0.99	0.92
II	0.3 ~ 0.5	65.4%	23.7%	2 – 4 – 10 – 12 – 20 – 21 – 23 – 30	0.99	0.93
III	0.5 ~ 0.7	70.9%	18.8%	1 – 4 – 6 – 10 – 12 – 15 – 17 - 23	0.99	0.95
IV	0.7 ~ 0.9	76.8%	12.3%	1 – 4 – 12 – 22 – 23 – 24 – 25 - 26	0.99	0.95
V	0.9 ~1.25	64.3%	24.3%	8 – 9 – 12 – 15 – 16 – 22 – 23 - 29	0.99	0.95
VI	1.25 ~ 2.0	69.8%	19.8%	5 – 7 – 13 – 15 – 19 – 23 – 28 - 31	0.99	0.92

4. APPLICATION OF THE PROPOSED SGMR SUBSETS FOR THE ANALYSIS OF A MDOF TEST STRUCTURE

The test structure, as shown in Figure 5, is a three-story asymmetric reinforced concrete frame, for which a pseudo-dynamic experiment was performed at full scale at the ELSA Laboratory [10]. The first mode period of the test structure (0.85 s) implies that subset IV of SGMR's, as described in Table 1, would be appropriate for the assessment of seismic collapse capacity. In order to evaluate the efficiency of the proposed subset records, the CDF of the mean value of the seismic collapse capacity based on the general SGMR set, as well as the proposed SGMR subsets (see Table 1), are shown in Figure 6a.

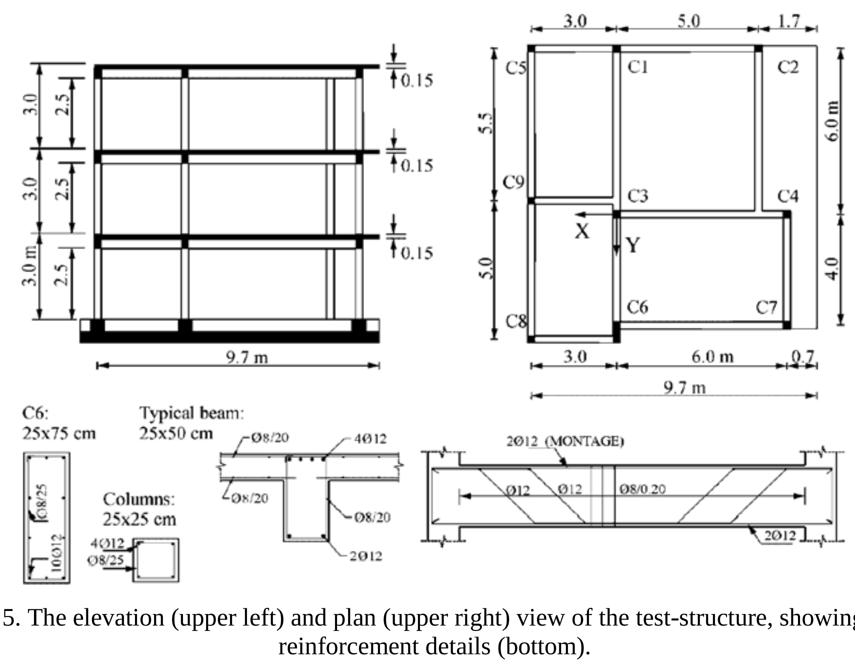

Figure 5. The elevation (upper left) and plan (upper right) view of the test-structure, showing typical reinforcement details (bottom).

Table 2. The SGMRs selected from the general set.

Event, Mw, Year	ID	Station, Dir	Vs_30 (m/s)	Campbell Distance (km)	Joyner-Boore Dist. (km)	PGA ('g)
Northridge, 6.7, 94	1	W Lost Cany, 000	309	12.4	11.4	0.41
	2	W Lost Cany, 270	309	12.4	11.4	0.48
Hector Mine, 7.1, 99	3	Hector, 000	685	12	10.4	0.27
Imperial Valley, 6.5, 79	4	Delta, 262	275	22.5	22	0.24
	5	Delta, 352	275	22.5	22	0.35
	6	El Centro Array #11, 140	196	13.5	12.5	0.36
	7	El Centro Array #11, 230	196	13.5	12.5	0.38
Kobe, Japan, 6.9, 95	8	Nishi-Akashi, 090	609	25.2	7.1	0.50
	9	Shin-Osaka, 000	256	28.5	19.1	0.24
Kocaeli, Turkey, 7.5, 99	10	Duzce, 180	276	15.4	13.6	0.31
	11	Duzce, 270	276	15.4	13.6	0.36
	12	Arcelik , 000	523	13.5	10.6	0.22
Landers, 7.3, 1992	13	Yermo Fire Station, 270	354	23.8	23.6	0.24
	14	Yermo Fire Station, 360	354	23.8	23.6	0.15
	15	Coolwater, LN	271	20	19.7	0.28
	16	Coolwater, TR	271	20	19.7	0.42
Loma Prieta, 6.9, 89	17	Capitola, 000	289	35.5	8.7	0.53
	18	Capitola, 090	289	35.5	8.7	0.44
	19	Gilroy Array #3, 000	350	12.8	12.2	0.56
Manjil, Iran, 7.4, 90	20	Abbar, T	724	13	12.6	0.50
Superstition Hills, 6.5, 87	21	El Centro Imp. Co. Cent, 000	192	18.5	18.2	0.36
	22	Poe Road (temp), 270	208	11.7	11.2	0.45
	23	Poe Road (temp), 360	208	11.7	11.2	0.30
Cape Mendocino, 7.0, 92	24	Rio Dell Overpass - FF, 270	312	14.3	7.9	0.39
	25	Rio Dell Overpass - FF, 360	312	14.3	7.9	0.55
Chi-Chi, Taiwan, 7.6, 99	26	CHY101, E	259	15.5	10	0.35
	27	TCU045, E	705	26.8	26	0.47
	28	TCU045, N	705	26.8	26	0.51
San Fernando, 6.6, 71	29	LA - Hollywood Stor FF, 090	316	25.9	22.8	0.21
	30	LA - Hollywood Stor FF, 180	316	25.9	22.8	0.17
Friuli, Italy, 6.5, 76	31	Tolmezzo, 000	425	15.8	15	0.35

The CDF for the SGMR subset IV, as shown in Figure 12a, shows good agreement with the general records ($S_1 = 0.99$, $S_2 = 0.95$), whereas the other subsets could not produce a good estimate for the mean of the seismic collapse capacity. Table 6 shows the mean values as well as the 99% confidence interval of the seismic collapse capacity according to each subset, and the associated similarity indices.

As an interesting result, the median IDA curve based on the general set and on each of the investigated SGMR subsets are shown in Figure 12b. Good agreement can be seen between the median IDA curve based on the general SGMR's and the median IDA curve obtained from subset IV, although the proposed methodology only works with the IM of the IDA curves corresponding to the collapse capacity points.

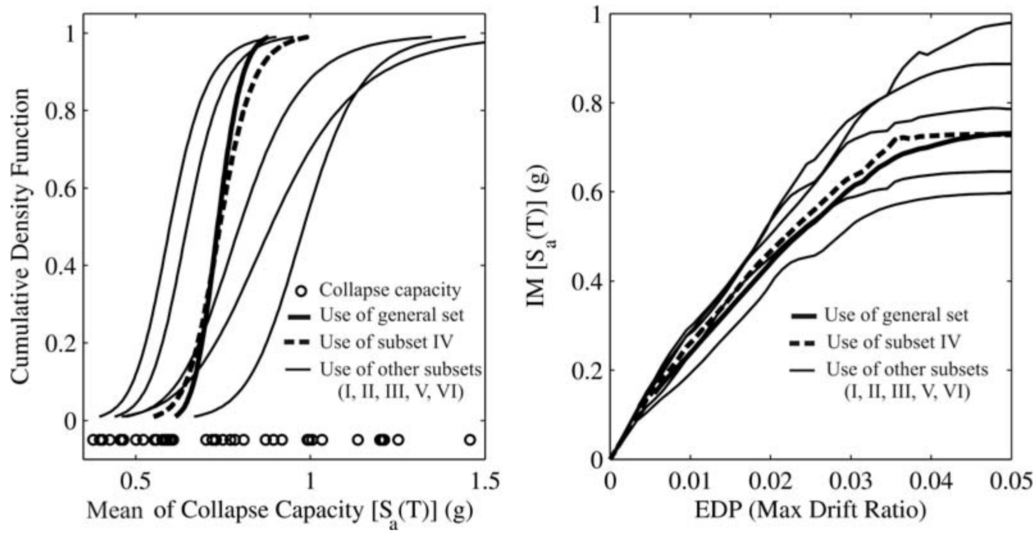

Figure 6. Efficiency assessment of the proposed subsets: (a) The CDF of the mean value of the seismic collapse capacity, (b) the mean of the full IDA analysis using different proposed subsets and the general set.

Table 3. The accuracy of using the proposed records for the seismic collapse capacity estimation of the test structure compared to the other subsets.

Ground Motion Class	Expected Value for Seismic collapse capacity	99% confidence Interval for Seismic collapse capacity	1^{st} Similarity Index	2nd Similarity Index
All Records	0.73	0.61 ~ 0.88	---	---
I	0.60	0.40 ~ 0.90	0.99	0.48
II	0.65	0.47 ~ 1.35	0.99	0.69
III	0.80	0.44 ~ 0.95	0.99	0.67
IV	0.74	0.55 ~ 1.00	1.00	0.92
V	0.98	0.67 ~ 1.45	0.98	0.26
VI	0.89	0.46 ~ 1.71	0.99	0.47

5. CONCLUSION

A new approach has been proposed in this paper for the categorization of SDOF systems into groups, and then a subset of SGMR's has been proposed for each of them. All of the presented subsets were chosen from a general set of SGMR's, and each of them is equivalent to the general set for the assessment of the seismic collapse capacity of the associated SDOF group. The grouping features are period and ductility. For this purpose, a statistical approach is first used in order to quantify the similarity of two sets of SGMR's for the assessment of the collapse capacity of an assumed SDOF group. Based on this approach, two similarity indices are defined, and GA is applied in order to determine the optimal subset of SGMR for the considered SDOF group. PC analysis is then used to explore the best pattern for grouping the SDOF systems. The best pattern for grouping the SDOF systems is the one that results in the minimum size of the corresponding SGMR subsets. The results of PC analysis have shown that the best selection can be achieved by grouping the SDOF systems, based on the period feature. Taking this into account, the range of the studied structural period values (0.1~2.0 second) has been divided into six groups, and eight records are proposed for each group (see Table 1). The results confirm the efficiency of the presented subsets of SGMR's for collapse capacity assessment. Also, the proposed subset shows good agreement with the general set for the prediction of the IM in the full range of EDP, i.e. maximum drift.

ACKNOWLEDGMENT
The research conducted by the authors has been funded by the International Institute of Earthquake Engineering and Seismology (IIEES) under Award Number 7134. This support is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the funding body.

REFERENCES

1. ASCE7-05, *Minimum Design Loads for Structures and other Structures*. American Society of Civil Engineers, Reston, VA, 2005.
2. Bommer J.J, Scott S.G, Sarma S.K. Hazard-consistent earthquake scenarios. *Soil Dynamics and Earthquake Engineering* 2000; 19: 219-231.
3. Naeim F, Lew M. On the use of design spectrum compatible time histories. *Earthquake Spectra* 1995; 11: 111-127.
4. Baker J.W, Cornell CA. Spectral shape, epsilon and record selection. *Earthquake Engineering and Structural Dynamics* 2006; **35**(9):1077–1095.
5. Baker J.W, Cornell CA. Vector-valued intensity measures incorporating spectral shape for prediction of structural response. *Journal of Earthquake Engineering* 2008; 12(4):534-554.
6. Haselton C.B, Baker J.W, Liel A.B, Deierlein G.G. Accounting for ground motion spectral shape characteristics in structural collapse assessment through an adjustment for epsilon. *ASCE Journal of Structural Engineering*; (in press).
7. Johnson AR, Wichern WD. *Applied multivariate statistical analysis*. Prentice Hall: 4th edition, 1998.
8. Goldberg DE. *Genetic Algorithms in search, Optimization, and Machine Learning*. Addison-Wesley: Reading, MA, 1989.
9. ATC 63, *Recommended Methodology for Quantification of Building System Performance and Response Parameters*, 90% Draft Report, Applied Technology Council, Redwood City, CA, 2008.
10. Negro P, Mola E, Molina FJ, Magonette GE. Full-scale testing of a torsionally unbalanced three-storey nonseismic RC frame. *Proceedings of the 13th World Conference on Earthquake Engineering*, Vancouver, Canada, 2004; 968.