

A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon

Jack W. Baker^{*,†} and C. Allin Cornell

Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, U.S.A.

SUMMARY

The ‘strength’ of an earthquake ground motion is often quantified by an Intensity Measure (*IM*), such as peak ground acceleration or spectral acceleration at a given period. This *IM* is used to predict the response of a structure. In this paper an intensity measure consisting of two parameters, spectral acceleration and epsilon, is considered. The *IM* is termed a vector-valued *IM*, as opposed to the single parameter, or scalar, *IMs* that are traditionally used. Epsilon (defined as a measure of the difference between the spectral acceleration of a record and the mean of a ground motion prediction equation at the given period) is found to have significant ability to predict structural response. It is shown that epsilon is an indicator of spectral shape, explaining why it is related to structural response. By incorporating this vector-valued *IM* with a vector-valued ground motion hazard, we can predict the mean annual frequency of exceeding a given value of maximum interstory drift ratio, or other such response measure. It is shown that neglecting the effect of epsilon when computing this drift hazard curve leads to conservative estimates of the response of the structure. These observations should perhaps affect record selection in the future. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: intensity measure; non-linear response; record selection; epsilon

INTRODUCTION

As non-linear dynamic analysis becomes a more frequently used procedure for evaluating the demand on a structure due to earthquakes, it is increasingly important to understand which properties of a recorded ground motion are most strongly related to the response caused in the structure. A value that quantifies the effect of a record on a structure is often called an Intensity Measure (*IM*). The Peak Ground Acceleration of a record was a commonly used *IM* in the past. More recently, spectral response values (e.g. spectral acceleration at the first-mode period of vibration – $S_a(T_1)$) have been used as *IMs*. Spectral acceleration at T_1 has

*Correspondence to: Jack W. Baker, Department of Civil and Environmental Engineering, Stanford University, MC 4020, Building 540, Stanford, CA 94305, U.S.A.

†E-mail: bakerjw@stanford.edu

Contract/grant sponsor: National Science Foundation (PEER); contract/grant number: EEC-9701568

Received 24 October 2004

Revised 28 January 2005

Accepted 2 February 2005

been found to be an effective *IM* [1], but among records with the same value of $S_a(T_1)$, there is still significant variability in the level of structural response in a multi-degree-of-freedom, non-linear structural model. If some of this remaining record-to-record variability could be accounted for by an improved intensity measure, then the accuracy and efficiency of structural response calculations could be improved.

In this paper, vector-valued intensity measures are considered as potential improvements on current intensity measures. The intensity measures considered consist of $S_a(T_1)$ as before, but also include a second parameter: the magnitude, distance or ε ('epsilon') associated with the ground motion. The *IM* is termed vector-valued because it now has two parameters, as opposed to traditional scalar, or single-parameter, *IMs*. It is found that the vector-valued *IM* consisting of $S_a(T_1)$ and ε is significantly superior to the *IM* consisting of $S_a(T_1)$ alone. The predictive power of ε is demonstrated, and an intuitive understanding is developed about the source of this predictive power.

WHAT IS EPSILON?

Magnitude and distance are familiar quantities to any earthquake engineer, but understanding of the ε parameter may be less common. Epsilon is defined by engineering seismologists studying ground motion as the number of standard deviations by which an observed logarithmic spectral acceleration differs from the mean logarithmic spectral acceleration of a ground-motion prediction (attenuation) equation. Epsilon is computed by subtracting the mean predicted $\ln S_a(T_1)$ from the record's $\ln S_a(T_1)$, and dividing by the logarithmic standard deviation (as estimated by the prediction equation). Epsilon is defined with respect to the unscaled record and will not change in value when the record is scaled. We will see later that ε is an indicator of the 'shape' of the response spectrum, and the shape of the spectrum does not change with scaling, providing intuition as to why ε would not vary with scaling.

Because of the normalization by the mean and standard deviation of the ground motion prediction equation, ε is a random variable with an expected value of zero, and a unit standard deviation. In fact, the distribution of ε is well represented by the standard normal distribution, at least within values of ± 3 [2, 3]. Thus, a sample of randomly chosen records will have an average ε value near zero, as can be seen visually in Figure 1(b), although an average value of zero is not required for our vector-valued *IM* work below. It should be noted that for a given ground motion record, ε is a function of T_1 (i.e. ε will have different values at different periods) and the ground motion prediction model used (because the mean and standard deviation of $\ln S_a(T_1)$ vary somewhat among models). The definition of ε is valid for any ground motion prediction model, but the model of Abrahamson and Silva [4] is the only one used in calculations here. If one would like to use ε in a vector *IM* to compute drift hazard, the model used to compute ε should be the same as the model used to perform the ground motion hazard assessment.

It should also be noted that there is more than one way to define spectral acceleration, and that the choice of definition will affect the computation of ε . Many ground motion prediction equations provide mean and standard deviation values for the average $\ln S_a(T_1)$ of the two horizontal components of a ground motion [4]. However, in this work we analyze only 2D frames, and thus use only one (arbitrarily chosen) component of a given ground motion. Thus, we use $\ln S_a(T_1)$ of an arbitrary horizontal component of the ground motion as our *IM*. It is

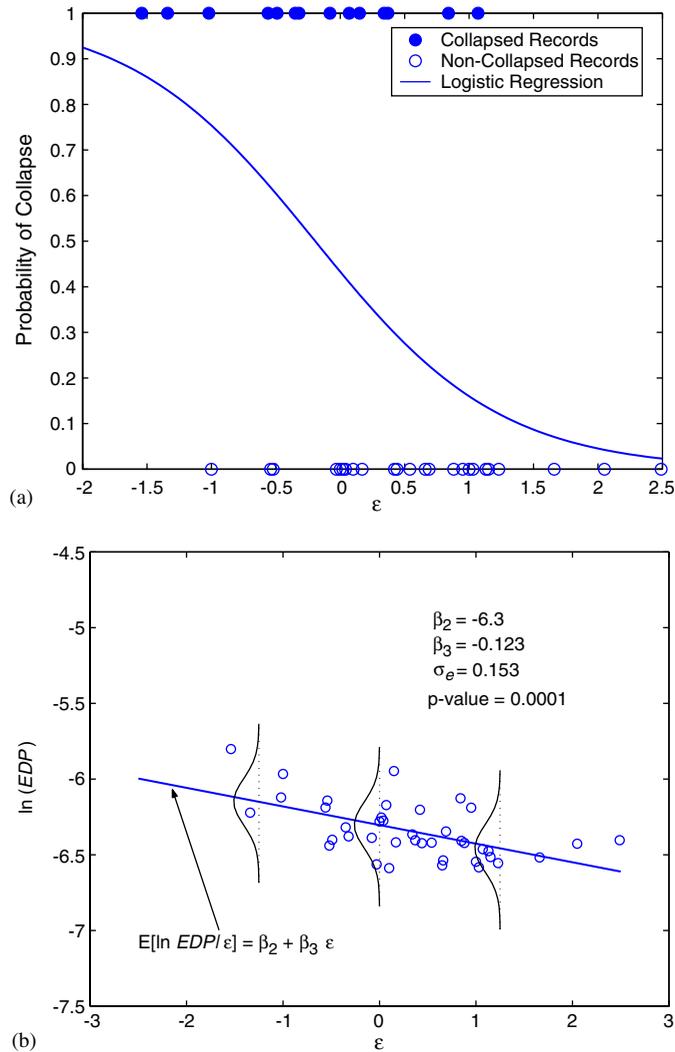


Figure 1. Analysis of data at a fixed value of S_a : (a) prediction of the probability of collapse using logistic regression applied to binary collapse/non-collapse results; and (b) prediction of response given no collapse, with the distribution of the residuals superimposed over the data.

therefore important that we compute our ground motion hazard for $\ln S_a(T_1)$ of an arbitrary component (as opposed to the average $\ln S_a(T_1)$ provided by the ground motion prediction equation). In this study, the ground motion hazard was computed using Abrahamson and Silva's equation, but the standard deviation of $\ln S_a(T_1)$ was inflated to reflect the increased variability of an arbitrary component of $\ln S_a(T_1)$ rather than the average of two components. The inflation factor was determined from another ground motion prediction model [5], which presents standard deviation values for both definitions of spectral acceleration. This inflated

standard deviation must also be used when computing the ε values for ground motion recordings as well. A more thorough explanation of this issue is in preparation by the authors.

Now that ε has been defined, we discuss how it and other *IMs* are used to predict drift, and then incorporated with ground motion hazard results to compute a drift hazard curve.

CALCULATION OF THE DRIFT HAZARD CURVE USING A SCALAR *IM*

Once an intensity measure is defined, the predicted structural response given an intensity measure level can be combined with Probabilistic Seismic Hazard Analysis (PSHA) to calculate the mean annual rate of exceeding a given structural response level. An example of the need for this calculation is seen in the work of the Pacific Earthquake Engineering Research (PEER) Center [6]. Here, following PEER practice, the response of a structure is termed an Engineering Demand Parameter, or *EDP* (in this paper, the only *EDP* considered is maximum interstory drift ratio, although the methodology is directly applicable to any *EDP* of interest). The annual frequency of exceeding a given level of the *EDP* is calculated as follows:

$$\begin{aligned}\lambda_{EDP}(z) &= \int_x P(EDP > z | IM = x) \cdot d\lambda_{IM}(x) \\ &\cong \sum_{\text{all } x_i} P(EDP > z | IM = x_i) \cdot \Delta\lambda_{IM}(x_i)\end{aligned}\quad (1)$$

where $\lambda_{EDP}(z)$ is the mean annual frequency of exceeding a given *EDP* value z , $\lambda_{IM}(x_i)$ is the mean annual frequency of exceeding a given *IM* value x_i (this is commonly referred to as the ground motion hazard curve), and $\Delta\lambda_{IM}(x_i) = \lambda_{IM}(x_i) - \lambda_{IM}(x_{i+1})$ is approximately the annual frequency of $IM = x_i$. The term $P(EDP > z | IM = x_i)$ represents the probability of exceeding a specified *EDP* level, z , given $IM = x_i$. In this paper we will use numerical integration to compute results, making use of the discrete summation approximation. We see that the rate of exceeding a given *EDP* level is found by assessing the ground motion hazard and the response of the structure, and coupling these two parts together with the use of an *IM*. Note that methods of computing drift hazard with techniques other than the *IM*-based method have been proposed [7, 8], but the *IM*-based method is the focus of this paper. If the *IM* is a vector, Equation (1) must be generalized, as will be discussed below.

PREDICTION OF STRUCTURAL RESPONSE USING A SCALAR *IM*

The *IM*-based procedure described above requires estimation of the probability distribution of structural response at a given *IM* level (i.e. $P(EDP > z | S_a(T_1) = x_i)$ in Equation (1)). An estimation procedure is now described for the scalar case, and later generalized to the case of a vector-valued *IM*. The scalar *IM* $S_a(T_1)$ is used in this section, both because of its wide use elsewhere, and because it will be easily generalized to our vector case: $S_a(T_1)$ and ε . Standard terminology from regression analysis is used in this section, to allow for quick descriptions of some concepts from statistics. The reader desiring a more detailed explanation is referred to a previous related publication [9].

The method used in this paper requires a suite of earthquake accelerograms, all at the same IM value, $S_a(T_1)=x$ (e.g. in this study, 40 records are used at each IM level). We scale a suite of recorded earthquake accelerograms to the given $S_a(T_1)$ value (e.g. Reference [1]). In this study, we use the same suite of records for different $S_a(T_1)$ levels, although one could use different record suites at different levels if PSHA disaggregation suggested that, for example, the representative magnitude level was changing [10]—we shall return to the record selection subject below. This suite of records is used to perform non-linear dynamic analysis on a model of the structure. Now we have n records, all with $S_a(T_1)=x$, and n corresponding values of EDP . So EDP given $S_a(T_1)=x$ is a random variable that we need to characterize.

Characterizing the collapses

When predicting non-linear response of structures, it is necessary to account for the possibility that some records may cause collapse of the structure at higher levels of IM . For the purpose of illustration here, collapse is defined to have occurred if the dynamic analysis algorithm fails to converge due to numerical instability or if the drift ratio at any story exceeds 10%. Such a finite cut-off is used because the validity of current non-linear models is not well confirmed beyond large deformation levels, which a real structure may not be able to reach before collapsing. For example, for the particular older reinforced concrete frame considered below, one might in fact expect axial failure of columns in the 3–5% interstory drift ratio range. This failure mechanism is difficult to model and was not incorporated in the computer model, resulting in larger displacements being obtained before collapse was signalled by numerical instability of the program. Other collapse criteria that have been used in such studies include dynamic instability, defined as the ratio of the increment in displacement to the increment in spectral acceleration level exceeding a specified threshold. The simpler 10% drift criterion is used here for illustration, but the proposed procedure applies universally, regardless of the structural model or the specified collapse criteria. Future research to more precisely model and identify response levels associated with collapse will be helpful to the drift hazard procedure presented here.

To account for these collapses, we separate our realizations of EDP into collapsed and non-collapsed data. We then estimate P , the probability of collapse, C , at the given $S_a(T_1)$ level. This estimate is denoted \hat{P} and calculated as:

$$\hat{P}(C | S_a(T_1)=x) = \frac{\text{number of records causing collapse}}{\text{total number of records}} \quad (2)$$

We then return to the non-collapse responses for the remainder of the response prediction.

Characterizing the non-collapse responses

The distribution of the non-collapsed responses for our EDP , maximum interstory drift ratio, has been found to be well represented by a lognormal distribution (the Kolmogorov–Smirnov test [11] was used to verify this supposition, and the same conclusion has been reached elsewhere [12, 13]). For this reason we work with the natural logarithm of EDP , which then follows the normal distribution. We can estimate the parameters for this normal distribution using the method of moments [14]. For each IM level, we denote the estimated mean of $\ln EDP$ as $\hat{\mu}_{\ln EDP | S_a(T_1)=x}$ and the estimated standard deviation as $\hat{\beta}_{\ln EDP | S_a(T_1)=x}$ (this

logarithmic standard deviation is sometimes referred to as ‘dispersion’). The probability that *EDP* exceeds z given $IM = x$ and no collapse can now be calculated using the normal complimentary cumulative distribution function:

$$P(EDP > z | S_a(T_1) = x, \text{ no collapse}) = 1 - \Phi \left(\frac{\ln z - \hat{\mu}_{\ln EDP | S_a(T_1) = x}}{\hat{\beta}_{\ln EDP | S_a(T_1) = x}} \right) \tag{3}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

Combining collapse and non-collapse results

We now combine the characterizations of collapses and non-collapse responses using the total probability theorem. Our estimate of the probability that *EDP* exceeds z given $IM = x$ is:

$$P(EDP > z | S_a(T_1) = x) = \hat{P}(C | S_a(T_1) = x) + (1 - \hat{P}(C | S_a(T_1) = x)) \left(1 - \Phi \left(\frac{\ln z - \hat{\mu}_{\ln EDP | S_a(T_1) = x}}{\hat{\beta}_{\ln EDP | S_a(T_1) = x}} \right) \right) \tag{4}$$

We can now proceed to work with this estimate.

CALCULATION OF THE DRIFT HAZARD CURVE USING A VECTOR-VALUED *IM*

A vector-valued *IM* can also be used to compute a drift hazard curve using a generalization of Equation (1), as given in Equation (5) [15].

$$\begin{aligned} \lambda_{EDP}(z) &= \int_{x_1} \int_{x_2} P(EDP > z | S_a(T_1) = x_1, \varepsilon = x_2) \cdot \left| \frac{\partial^2 \lambda_{IM}(x_1, x_2)}{\partial x_1 \partial x_2} \right| dx_1 dx_2 \\ &\cong \sum_{\text{all } x_{1,i}} \sum_{\text{all } x_{2,j}} P(EDP > z | S_a(T_1) = x_{1,i}, \varepsilon = x_{2,j}) \cdot \Delta \lambda_{IM}(x_{1,i}, x_{2,j}) \end{aligned} \tag{5}$$

We see first that the scalar-*IM* drift prediction $P(EDP > z | IM = x)$ has been replaced with the vector-*IM* drift prediction $P(EDP > z | S_a(T_1) = x_1, \varepsilon = x_2)$, which will be expanded by a means analogous to Equation (4) in Equation (11) below. In addition, the scalar-*IM* ground motion hazard has been replaced by the joint hazard of the vector-valued *IM*. Defining $\Delta \lambda_{IM}(x_{1,i}, x_{2,j})$ as $\lambda_{S_a \in [x_{1,i}, x_{1,i+1}], \varepsilon \in [x_{2,j}, x_{2,j+1}]}$, we take advantage of the fact that we could also express this as the marginal rate density of $S_a(T_1)$, and the conditional probability distribution of ε given $S_a(T_1)$:

$$\Delta \lambda_{IM}(x_{1,i}, x_{2,j}) = P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) = x_{1,i}) \cdot \Delta \lambda_{S_a(T_1)}(x_{1,i}) \tag{6}$$

Then Equation (5) can be restated as:

$$\begin{aligned} \lambda_{EDP}(z) &= \sum_{\text{all } x_{1,i}} \sum_{\text{all } x_{2,j}} P(EDP > z | S_a(T_1) = x_{1,i}, \varepsilon = x_{2,j}) \\ &\quad \times P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) = x_{1,i}) \cdot \Delta \lambda_{S_a(T_1)}(x_{1,i}) \end{aligned} \tag{7}$$

We state the equation in this way because we obtain the distribution of $S_a(T_1)$ and ε from PSHA in this form: $\Delta\lambda_{S_a(T_1)}(x_{1,i})$ comes from the standard PSHA hazard curve, and $P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) = x_{1,i})$ is a standard disaggregation result. Note that the disaggregation can be presented in more than one way. Some PSHA codes provide $P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) = x_{1,i})$ (see Reference [16]), while others provide $P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) \geq x_{1,i})$ (see Reference [17]). For instance, Abrahamson's code [18] provides $P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) = x_{1,i})$, while the U.S. Geological Survey [19] provides $P(x_{2,j} < \varepsilon < x_{2,j+1} | S_a(T_1) \geq x_{1,i})$. It is fairly simple matter to convert the results between the two forms [20, p. 195], but one should be aware of which version is provided by the software in use, and convert the results if necessary. The hazard assessments in this study were performed using the Abrahamson code, which provides results (e.g. Figure 6) directly in the form needed for drift hazard calculations.

PREDICTION OF BUILDING RESPONSE USING A VECTOR-VALUED *IM*

We now generalize the prediction procedure of the preceding section for use with a vector-valued *IM*. Consider the vector consisting of $S_a(T_1)$ and ε . We are now trying to estimate $P(EDP > z | S_a(T_1) = x_1, \varepsilon = x_2)$. The simplest solution, if possible, would be to scale our records to both parameters of our *IM*: $S_a(T_1) = x_1$ and $\varepsilon = x_2$. However, ε is defined with respect to the unscaled record, and does not change with scaling (similarly, magnitude and distance do not change with scaling). Because of this, we need a supplement to scaling for the vector-valued *IM* procedure, in order to predict response as a function of the second *IM* parameter.

The solution we adopt is to scale to $S_a(T_1)$ as before, and then apply regression analysis to estimate *EDP* as a function of ε [11]. Thus our treatment of $S_a(T_1)$ remains the same as in the scalar case, but we now incorporate information from the regression on ε . Our approach with the vector case is the same as the scalar case in that we separate out the collapse responses first, and then deal with the remaining non-collapse responses.

Accounting for collapses with the vector-valued IM

When using a vector-valued *IM*, instead of taking the probability of collapse to be simply the fraction of records that cause collapse, we can take advantage of the second *IM* parameter to predict the probability of collapse more accurately. We do this using logistic regression, which is a commonly used tool for analyzing binary data [11]. It should be noted that this is not the only method for quantifying the probability of collapse. For example, a bivariate normal model for collapse *capacity* of the structure could be defined and used to estimate probability of collapse. The results from the two models will be nearly identical, but it may be more convenient to adopt one or the other in certain cases. This will be explored in a future publication by the authors.

With the logistic regression procedure used here, each record has a value of ε , which we use as our predictor variable. We designate C as an indicator variable for collapse (C is equal to 1 if the record causes collapse and 0 otherwise). We then use the logistic regression to

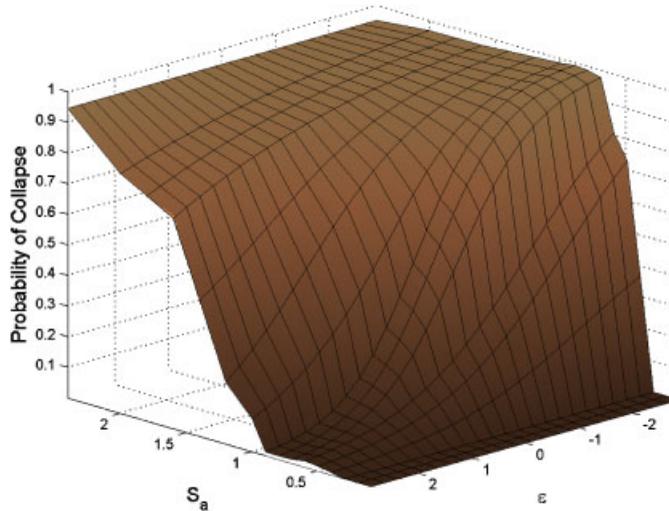


Figure 2. Prediction of the probability of collapse as a function of both $S_a(T_1)$ and ε .

predict collapse:

$$\hat{P}(C | S_a(T_1) = x_1, \varepsilon = x_2) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_2)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_2)} \quad (8)$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are coefficients to be estimated from regression on a dataset that has been scaled to $S_a(T_1) = x_1$ (i.e. $\hat{\beta}_0$ and $\hat{\beta}_1$ will be different for different values of $S_a(T_1)$). An example of this data and a fitted logistic regression curve is presented in Figure 1(a). The tendency for the probability of collapse to decay with increasing ε is common; such observations will be discussed below. By performing this regression for all $S_a(T_1)$ levels, one can obtain the probability of collapse as a function of both $S_a(T_1)$ and ε , as seen in Figure 2.

It should be noted that the level of confidence in the result from Equation (8) depends on the nature of the data used in the regression analysis. If there are significant numbers of both collapses and non-collapses in the dataset, then the regression should be very stable. However, if for a given S_a level, the dataset consists of, for instance, 39 records that do not cause collapse, and one record that causes collapse, the logistic regression will be strongly influenced by the single collapse data point, and may indicate a different trend that in fact exists (if the exercise were to be repeated with more records or different records). For this reason, it is suggested that if there are two or fewer collapse data points, then the probability of collapse should be taken as a simple constant (i.e. $1/n$ or $2/n$, where n is the number of records) for all levels of x_2 . This is equivalent to using the scalar-*IM* procedure for the collapse portion of the prediction (Equation (2)). The same should be done in the case where all but one or two records cause collapse. It should be noted that in these cases, the probability of collapse will already be very high or very low, so the second *IM* parameter would not have much effect anyway. This modification to the procedure will merely prevent the logistic regression prediction from producing unstable results.

Characterizing non-collapses with the vector-valued IM

We now incorporate the second parameter of our *IM* in prediction of response for the non-collapse results. Note that each of these records has been scaled to $S_a(T_1) = x_1$. Each of the records has a value of ε and a value of *EDP*. We have found that there tends to be a relationship between ε and *EDP* of the form $\ln EDP = \beta_2 + \beta_3\varepsilon + e$, where β_2 and β_3 are constant coefficients, and e is the prediction error ('residual'). We can use linear least-squares regression [11] to obtain estimates of the two regression coefficients, $\hat{\beta}_2$ and $\hat{\beta}_3$ (again, these values will vary for different $S_a(T_1)$ levels). A graphical example of this data and the regression fit is shown in Figure 1(b).

When using linear least-squares regression on a dataset, several assumptions are normally implicitly made, and the accuracy of the results depends on the validity of these assumptions. The prediction error of record i (the difference between the predicted value of $\ln EDP_i$ and the actual value) is termed the 'residual' of record i , and is assumed to be mutually independent from the prediction error of record j for all $i \neq j$. In addition, we will later assume the residuals to be normally distributed with constant variance (i.e. homoscedastic). The assumptions of independent normal residuals with constant variance have been examined for the data in this study, and found to be reasonable. An estimate of the variance of the residuals is also available from the analysis software, and we denote it $\text{Var}[e] \equiv \hat{\sigma}_e^2$. This variance in the residuals is displayed graphically in Figure 1(b), by superimposing the estimated normal distribution of the residuals over the data. From regression, we now know that given $S_a(T_1) = x_1$, $\varepsilon = x_2$ and no collapse, the mean value of $\ln EDP$ is:

$$E[\ln EDP] = \hat{\beta}_2 + \hat{\beta}_3 x_2 \tag{9}$$

where $\hat{\beta}_2$ and $\hat{\beta}_3$ have been obtained by regressing on records scaled to $S_a(T_1) = x_1$. We also know that, conditional on $S_a(T_1)$ and ε , $\ln EDP$ is normally distributed with variance equal to the $\hat{\sigma}_e^2$. So the probability that *EDP* exceeds z , given $S_a(T_1) = x_1$, $\varepsilon = x_2$, and no collapse can be expressed as:

$$P(EDP > z | S_a(T_1) = x_1, \varepsilon = x_2, \text{no collapse}) = 1 - \Phi \left(\frac{\ln z - (\hat{\beta}_2 + \hat{\beta}_3 x_2)}{\hat{\sigma}_e} \right) \tag{10}$$

Recall that $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\sigma}_e^2$ are all functions of $S_a(T_1)$ or x_1 . This equation is very similar to Equation (3) used in the scalar case. We previously estimated the mean of the normal distribution by the average logarithmic response of all records, but now we use a result from regression on ε . We have also replaced the standard deviation of the records by the standard deviation of the regression residual. Otherwise, the equation is the same.

As with the estimation of collapse, there is more than one way to incorporate the vector *IM*. For example, rather than using regression, one could use a weighted scheme where the weights of each record are determined according to the results of the PSHA analysis [12]. The results from these two schemes will agree closely, but one may be more appropriate than the other based on the amount of data available and the extent to which the effect of the *IM* can be parameterized. This will be described in a future publication by the authors.

We now combine the possibilities of collapse or no collapse, using the Total Probability Theorem and Equations (8) and (10), to compute the conditional probability that *EDP*

exceeds z :

$$P(EDP > z | S_a(T_1) = x_1, \varepsilon = x_2) = \hat{P}(C) + (1 - \hat{P}(C)) \left(1 - \Phi \left(\frac{\ln z - (\hat{\beta}_2 + \hat{\beta}_3 x_2)}{\hat{\sigma}_e} \right) \right) \quad (11)$$

where

$$\hat{P}(C) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_2)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_2)}$$

Although x_1 does not appear in Equation (11), our estimate is implicitly a function of x_1 , because the data used to estimate $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\sigma}_e$ all come from records scaled to $S_a(T_1) = x_1$. This gives us a response prediction that is similar to the original prediction of Equation (4), but that now incorporates a two-element vector.

INVESTIGATION OF MAGNITUDE, DISTANCE AND EPSILON AS *IM* PARAMETERS

The procedure for evaluating drift hazard using a vector-valued *IM* can now be used to assess the response-predicting effectiveness of ε as an element in a vector with $S_a(T_1)$. Additionally, we will examine magnitude (M) and distance (R) as elements in a vector with $S_a(T_1)$ using the same procedure, to evaluate whether they have any significant effect on structural response after conditioning on $S_a(T_1)$. This study considers only M , R and ε because the conditional distribution $P(x_{2,j} < IM_2 < x_{2,j+1} | S_a(T_1) = x_{1,i})$ is then easily available from standard PSHA software (where IM_2 is used to represent either M , R or ε). If one is interested in the effect of other parameters such as spectral shape or duration, special modifications to the PSHA analysis [15] are needed in order to obtain the conditional distribution $P(x_{2,j} < IM_2 < x_{2,j+1} | S_a(T_1) = x_{1,i})$.

To test the effectiveness of M , R or ε as predictors, we examine the results from our response regressions on these variables. The seven storey concrete frame structure described below was used to generate response data. At each level of $S_a(T_1)$, we predict collapse using logistic regression on M , R or ε , and we use linear regression on these variables to model the non-collapse responses. An effective predictor should show a trend in one or both of these regressions and the trend (i.e. slope of the regression) should be statistically significant. A standard way of measuring statistical significance is with the 'p-value' for the regression coefficient. Typically, a p-value smaller than 0.05 is interpreted to indicate that the predictive variable is significant [11], although when a slightly larger p-value shows up repeatedly in separate tests of the same predictor variable, this can also be interpreted as an indicator of significance. A p-value can also be computed for both the logistic and linear regression results.

When distance is considered as a candidate *IM* parameter, we find no statistical significance (median p-values of 0.34 and 0.29 for linear and logistic regression, respectively, considering 13 levels of $S_a(T_1)$). Both ε and magnitude show some significance for the linear response regression (median p-values of 0.05 and 0.06, respectively), although ε shows more significance than M for the logistic collapse probability regression (median p-values of 0.14 and

Table I. *P*-values from linear and logistic regression on magnitude and epsilon.

S_a	Percent collapsed	Magnitude			Potential Prediction error ($M = 6.5$ vs $M = 7$)
		Linear regression <i>p</i> -value	Logistic regression <i>p</i> -value	Linear regression coefficient	
0.1	0%	0.04	–	0.13	7%
0.2	0%	0.09	–	0.16	8%
0.3	0%	0.05	–	0.25	13%
0.4	0%	0.04	–	0.29	16%
0.5	0%	0.06	–	0.35	19%
0.6	3%	0.06	–	0.37	20%
0.7	8%	0.05	0.36	0.33	18%
0.8	9%	0.20	0.45	0.23	12%
0.9	14%	0.10	0.29	0.36	20%
1.0	16%	0.04	0.37	0.41	22%
1.2	24%	0.10	0.10	0.28	15%
1.4	56%	0.10	0.98	0.49	28%
1.6	68%	0.51	0.66	0.27	14%
Median		0.06	0.37	0.29	16%
S_a	Percent collapsed	Epsilon			Potential Prediction error ($\varepsilon = 0$ vs $\varepsilon = 1.5$)
		Linear regression <i>p</i> -value	Logistic regression <i>p</i> -value	Linear regression coefficient	
0.1	0%	0.00	–	–0.15	24%
0.2	0%	0.01	–	–0.15	25%
0.3	0%	0.07	–	–0.14	24%
0.4	0%	0.05	–	–0.19	32%
0.5	0%	0.01	–	–0.30	56%
0.6	3%	0.13	–	–0.20	34%
0.7	8%	0.03	0.14	–0.24	42%
0.8	9%	0.01	0.26	–0.30	57%
0.9	14%	0.02	0.20	–0.32	61%
1.0	16%	0.08	0.14	–0.24	43%
1.2	24%	0.13	0.04	–0.19	32%
1.4	56%	0.17	0.09	–0.23	42%
1.6	68%	0.69	0.05	–0.09	14%
Median		0.05	0.14	–0.20	34%

0.37, respectively). The lack of significance of R and slight significance of M are consistent with results from previous work on this topic (e.g. Reference [1]). The significance of ε and the evaluation of significance with respect to collapse prediction are believed to be new results.

The *p*-values from both regressions using M and ε as predictors at 13 levels of spectral acceleration are given in Table I (distance has been omitted because of the consistent lack of significance it demonstrated). Values for logistic regression are omitted when fewer than

three records cause collapse, preventing the regression from being performed. At high $S_a(T_1)$ levels, when nearly all of the records cause collapse and there is little data for either the linear or logistic regression, the regressions are less useful, and the results are omitted.

In addition to standard p-values, a value termed the 'potential prediction error' is also reported in Table I. This value is defined as the percentage difference in predicted response for a reasonable range of values in the second *IM* (i.e. what is the difference in response between magnitude 6.5 and 7 records, given the same $S_a(T_1)$ level). In Table I we see that for the given range of variation, a change in magnitude produces a potential prediction error of 16%, while a change in ε produces a potential prediction error of 34%. This major difference is because, while the slopes of the two trends are similar, there is more room for variation with ε (records are typically selected to be within 0.5 magnitude units of the target value obtained from PSHA, but using zero-epsilon records in place of 1.5-epsilon records is not uncommon). Both the statistical significance and this potential prediction error are relevant, and jointly they suggest that ε , and to a lesser extent magnitude, should be considered when predicting the response of a structure.

Note that the trend with ε would be more difficult to discern if we had not already scaled the records to $S_a(T_1)$. This is because ε and $S_a(T_1)$ tend to be correlated, making it more difficult to separate effects due to $S_a(T_1)$ and effects due to ε . By first scaling to $S_a(T_1)$ we have eliminated this problem (termed 'collinearity' [11]) and allowed the effect of ε to be seen more clearly.

Although Table I appears to show empirically that ε has an effect on structural response, this conclusion would be much more convincing if supported by an intuitive understanding as to why ε might matter. This understanding is developed in the following section.

WHY DOES EPSILON AFFECT STRUCTURAL RESPONSE?

When considering why ε could affect structural response, we should consider current understanding about non-linear response of multiple-degree-of-freedom structures. The first parameter of our *IM*, $S_a(T_1)$, provides the response of a linear single-degree-of-freedom structure with a period of vibration approximately equal to the first-mode period of the MDOF structure under consideration. Given $S_a(T_1)$, the shape of the response spectra is known to be a significant factor in the response of non-linear MDOF structures (e.g. References [9, 21]). This is because the response of an MDOF structure is also affected by excitation of higher modes of the structure at periods shorter than T_1 (as implied, e.g. by response spectrum analysis [22]). In addition, S_a at periods longer than T_1 affects non-linear structures because, as the structure starts behaving non-linearly, the effective period of its first mode increases to a period larger than T_1 [23, 24]. Thus, given two records with the same $S_a(T_1)$ value, the record with higher S_a values at periods other than T_1 will tend to cause larger responses in a non-linear MDOF system. So given $S_a(T_1)$, we would like to know about S_a at other periods. We could do this by measuring that S_a at other periods directly [9, 25, 26]. Alternatively, we will see that ε is a convenient implicit measure of spectral shape.

A record with a positive ε value is one that has a larger-than-expected spectral acceleration at the specified period. But what does it tell us about the spectral acceleration at other periods? It may be that the record is stronger than expected at all periods, or it may be that the record is stronger than expected in only a nearby range of periods, and that the spectral acceleration

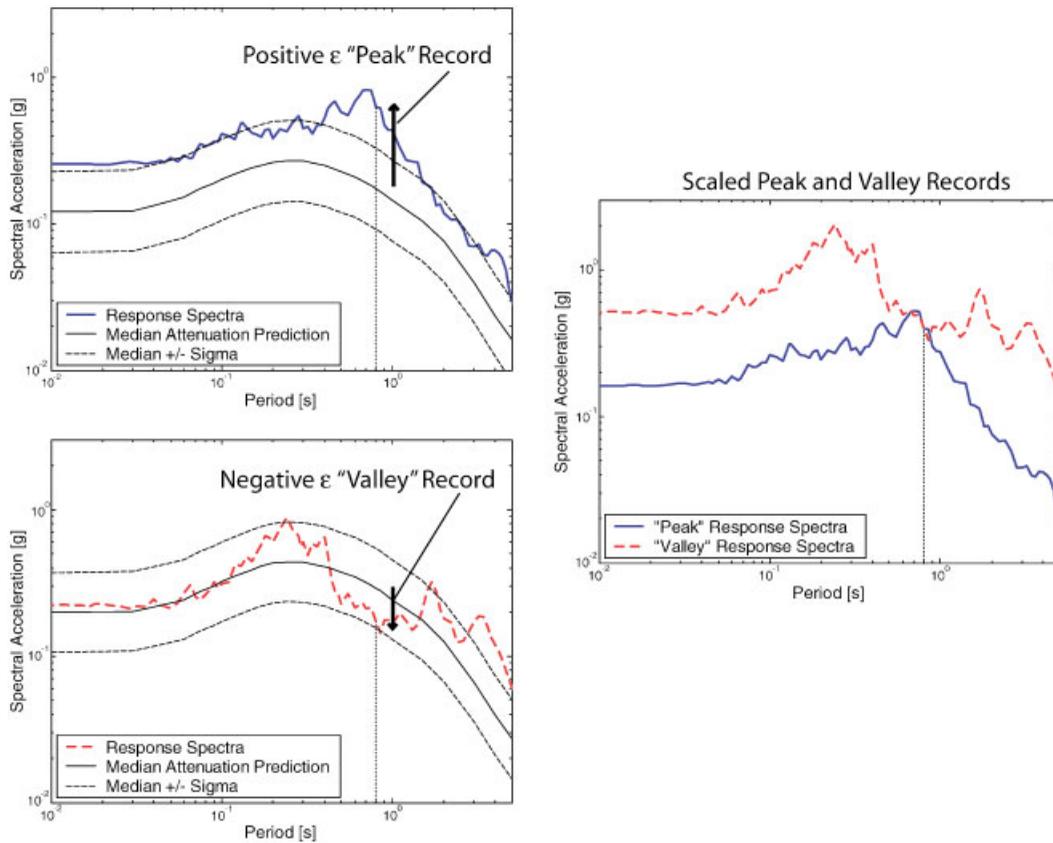


Figure 3. Scaling a negative ε record and a positive ε record to the same $S_a(T_1)$: an illustration of the peak and valley effect. In this case, $T_1 = 0.8$ s.

values at other periods are not as strongly related. We are interested in the possibility that only a narrow range of periods have comparatively large S_a values. We term this a record with a spectral 'Peak' at $S_a(T_1)$. Conversely, a record that is lower than expected in only a narrow range of periods has a spectral 'Valley' at $S_a(T_1)$. Now consider scaling a record with a peak and a record with a valley to the same $S_a(T_1)$ level. At T_1 , the two records will have the same spectral acceleration by construction, but at other periods the valley record will tend to have larger spectral accelerations than the peak record. This is seen by examining the two sample response spectra shown in Figure 3. If a record has a peak or a valley at the period considered, then ε (which measures deviation from expected spectral values) may be an indicator of this condition, and if so it would be useful for predicting structural response.

The effect of epsilon, as seen using a second-moment model for logarithmic spectral acceleration

Anecdotal evidence of ε indicating a spectral peak or valley is seen in Figure 3, but there is more concrete evidence to show the connection between ε and spectral shape. We note

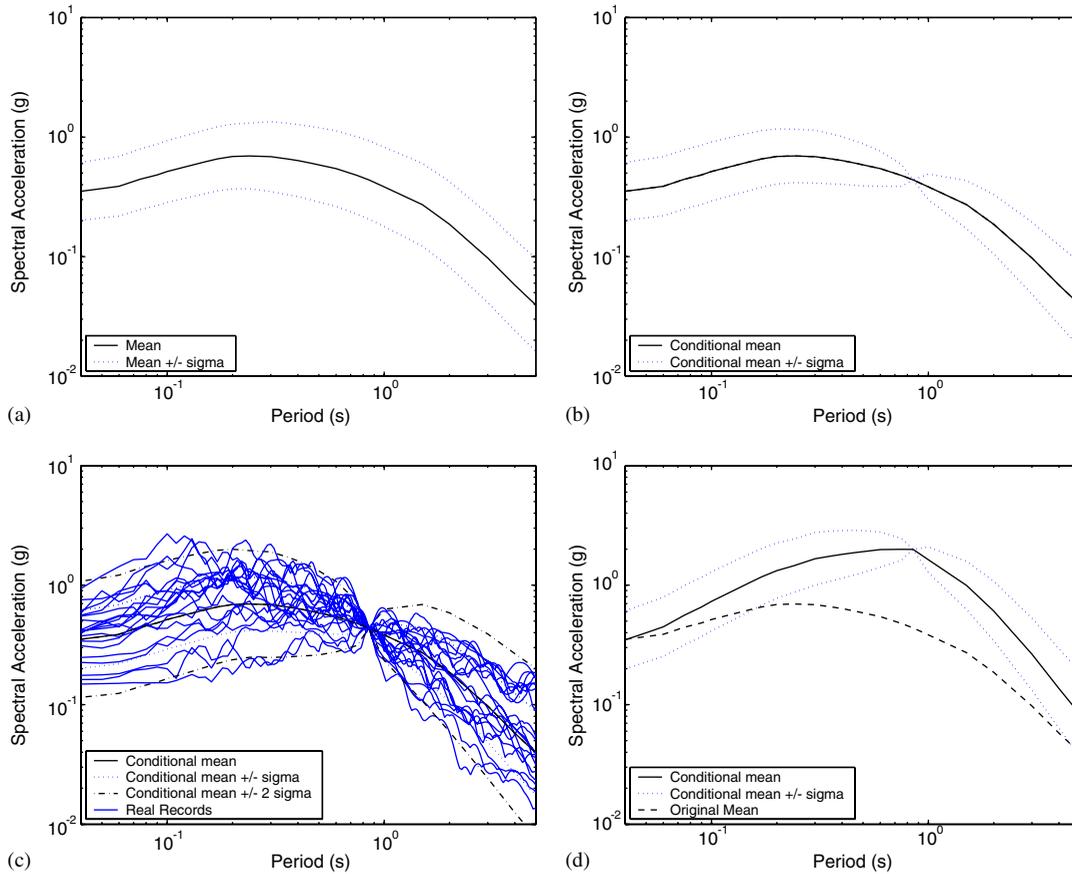


Figure 4. The mean value and the mean \pm sigma values of $\ln S_a$ for a Magnitude = 6.5, Distance = 8 km event: (a) unconditioned values; (b) conditioned on $\ln S_a(0.8 \text{ s})$ equal to the mean value of the ground motion prediction equation; (c) conditioned on $\ln S_a(0.8 \text{ s})$ equal to the mean value of the ground motion prediction with actual response spectra superimposed; and (d) conditioned on $\ln S_a(0.8 \text{ s})$ equal to the mean value of the ground motion prediction plus two standard deviations.

that for a given magnitude, distance, site classification and faulting mechanism, logarithmic spectral acceleration at a given period is a random variable with a mean and standard deviation specified by a ground motion prediction equation [4]. Lines indicating the mean value \pm one standard deviation at all periods for a scenario event are shown in Figure 4(a). Now consider logarithmic spectral accelerations at two periods simultaneously. We can obtain the means and standard deviations from the ground motion prediction equation. In order to completely specify the first and second moments of this pair, we also need to know the correlation between $\ln S_a$ values at the two periods. An empirically determined relationship for this correlation is given by Inoue and Cornell [27]:

$$\rho_{\ln S_a(T_1), \ln S_a(T_2)} = 1 - 0.33 |\ln(T_1/T_2)| \quad 0.1 \text{ s} \leq T_1, T_2 \leq 4 \text{ s} \quad (12)$$

This correlation coefficient approaches one when the two periods are nearly equal, and decreases as the periods get further apart from each other. We have now fully defined the mean and covariance of this pair of response spectra values. Note that for expository purposes, we will later assume that this equation is valid over the range $0.05 \text{ s} \leq T_1, T_2 \leq 5 \text{ s}$.

Previous research has established that $\ln S_a(T_1)$ and $\ln S_a(T_2)$ are each marginally normally distributed. Under the mild assumption that they are jointly normally distributed, we obtain the conditional mean of $\ln S_a(T_2)$, given $\ln S_a(T_1)$, as given in Equation (13):

$$\begin{aligned} \mu_{\ln S_a(T_2) | \ln S_a(T_1) = x} &= \mu_{\ln S_a(T_2)} + \rho_{\ln S_a(T_1), \ln S_a(T_2)} \cdot \sigma_{\ln S_a(T_2)} \left(\frac{x - \mu_{\ln S_a(T_1)}}{\sigma_{\ln S_a(T_1)}} \right) \\ &= \mu_{\ln S_a(T_2)} + \rho_{\ln S_a(T_1), \ln S_a(T_2)} \cdot \sigma_{\ln S_a(T_2)} \cdot \varepsilon_{T_1} \end{aligned} \tag{13}$$

where ε_{T_1} is the ε value of the record at the period T_1 (this equation is derived from the result $E[\varepsilon_{T_2} | \varepsilon_{T_1}] = \rho_{\ln S_a(T_1), \ln S_a(T_2)} \cdot \varepsilon_{T_1}$). We see that the conditional mean of $\ln S_a(T_2)$ is shifted up if $\varepsilon_{T_1} > 0$ or down if $\varepsilon_{T_1} < 0$. We can also obtain the conditional standard deviation of $\ln S_a(T_2)$:

$$\sigma_{\ln S_a(T_2) | \ln S_a(T_1) = x} = \sigma_{\ln S_a(T_2)} \sqrt{1 - \rho_{\ln S_a(T_1), \ln S_a(T_2)}^2} \tag{14}$$

We see that, as might be expected, the conditional standard deviation of $\ln S_a(T_2)$ is reduced as the correlation increases between $\ln S_a(T_1)$ and $\ln S_a(T_2)$. In Figure 4(b), for $T_1 = 0.8 \text{ s}$, we condition on $\ln S_a(0.8 \text{ s}) = \mu_{\ln S_a(T_1)}$ and plot the mean and mean \pm sigma for a scenario event (for each value of T , we use the conditional mean and standard deviation from Equations (13) and (14)). This plot represents the theoretical distribution of the response spectra for all records with $\ln S_a(0.8 \text{ s}) = \mu_{\ln S_a(T_1)}$.

We see in Figure 4(b) that there is less dispersion in S_a for periods close to T_1 , but for periods at some distance from T_1 there is little reduction in dispersion gained by conditioning on T_1 . While it is not strictly a verification, we can confirm that this multivariate distribution model is representative of reality by scaling real records to $\ln S_a(0.8 \text{ s})$ and superimposing them over the model distribution. We see in Figure 4(c) that the real records match the model reasonably well (i.e. the mean value of the records is close to the predicted mean value, and the model prediction that 95% of the records should fall between the mean \pm two sigma is reasonable).

In Figure 4(d) we plot the expected S_a value and \pm sigma, but now conditioned on $\ln S_a(0.8 \text{ s}) = \mu_{\ln S_a(T_1)} + 2\sigma_{\ln S_a(T_1)}$ (i.e. an ‘ $\varepsilon = 2$ ’ record). The original mean is also plotted as a reference. In this figure, we see that for periods near to T_1 we have larger spectral values than originally expected, but as the period gets much larger or smaller than T_1 , the expected value of $\ln S_a(T)$ goes back towards the original mean value, reflecting Equations (12) and (13).

The results of this analysis can be restated in words as follows. A record with a positive ε has a higher than expected S_a value at the specified period. But S_a values are not perfectly correlated, so a higher-than-average value at one period does not imply correspondingly higher-than-average values at all periods—in fact, the conditional expected values of S_a at other periods tend back towards the marginal expected value. Thus, records with positive ε values tend to have peaks in the response spectrum at the specified period, and records with negative ε values tend to have valleys. Therefore, ε is an indicator of spectral shape, and this is why it is effective in predicting the response of non-linear MDOF models.

Consideration of other candidate IM parameters

It has been shown that ε is an indicator of spectral shape, supporting its utility in predicting non-linear response. But the question naturally arises, is it the *best* predictor of non-linear response? Other candidates have been considered. Magnitude has often been considered as an indicator of spectral shape as discussed above, but it effectively has a weaker relationship with spectral shape than ε . We demonstrate next via Figure 5 this difference. Consider a target event with $M = 7$ and $\varepsilon = 1.5$. Records with these parameters are rare; there are few recordings available from large magnitudes, and of those recordings, only 7% are expected to have an $\ln S_a(T_1)$ at least 1.5 standard deviations larger than the mean. Suppose instead we have two records available for analysis: an $M = 7$, $\varepsilon = 0$ record and an $M = 6.5$, $\varepsilon = 1.5$ record (i.e. we can match either magnitude or ε , but not both). The expected spectra for these two records are shown in Figure 5(a), along with the expected spectrum for the target event. In Figure 5(b), the three 'records' have been scaled so that their $S_a(T_1)$ values match. We see that the record with the correct ε and incorrect magnitude comes closer to matching the target spectral shape than does the record with the correct magnitude but incorrect ε . So a difference of 0.5 units in magnitude makes much less difference in spectral shape than the difference of 1.5 units of ε . Therefore it is anticipated that ε will prove more effective than M . This conclusion will be confirmed below.

As mentioned earlier, it is also possible to measure spectral shape in a direct manner by considering spectral acceleration at an additional period as a second parameter in the vector *IM*. In previous work, the drift estimation method presented in this paper has been used with such a direct measure of spectral shape. With an informed choice of the second period, the direct measure of spectral shape has been shown to be an effective *IM* as well [9]. The comparative advantage of ε is that it is easy to find the joint distribution of S_a and ε from PSHA disaggregation, with no need for specialized hazard analysis software. Further consideration of spectral shape is outside of the scope of this paper, but it should also be considered a promising candidate for a vector-valued *IM*.

Spectral acceleration averaged over a range of periods may be an effective *IM* in the case where a structure is sensitive to several periods, and may reduce the peak/valley effects seen when using spectral acceleration at only a single period [12]. The disadvantage is that when structural response is predominantly governed by a single period, averaging spectral values over a range of periods will tend to reduce the efficiency of response predictions.

One might also wonder if there is a better epsilon-based measure of spectral shape. For example, it is possible that a record could have a positive ε at the period considered, but also have large ε values at all other periods. In this case, the positive ε value would incorrectly suggest the presence of a peak and thus it would incorrectly predict the relatively smaller level of response we have seen to be typically associated with positive ε values. It would seem plausible that a more sophisticated 'epsilon-type' parameter might do a better job than one considering simply ε itself. The authors have tried several approaches to develop an improved epsilon-based measure. A measure separating the inter-event epsilon from the intra-event epsilon (see Reference [4]) was attempted, anticipating that the intra-event epsilon would tend to represent peaks and valleys, while the inter-event epsilon would reflect the overall increase or decrease in the response spectra already accounted for by $S_a(T_1)$. Measures attempting to separate 'global' shape effects from 'local' shape effects (see Reference [28])

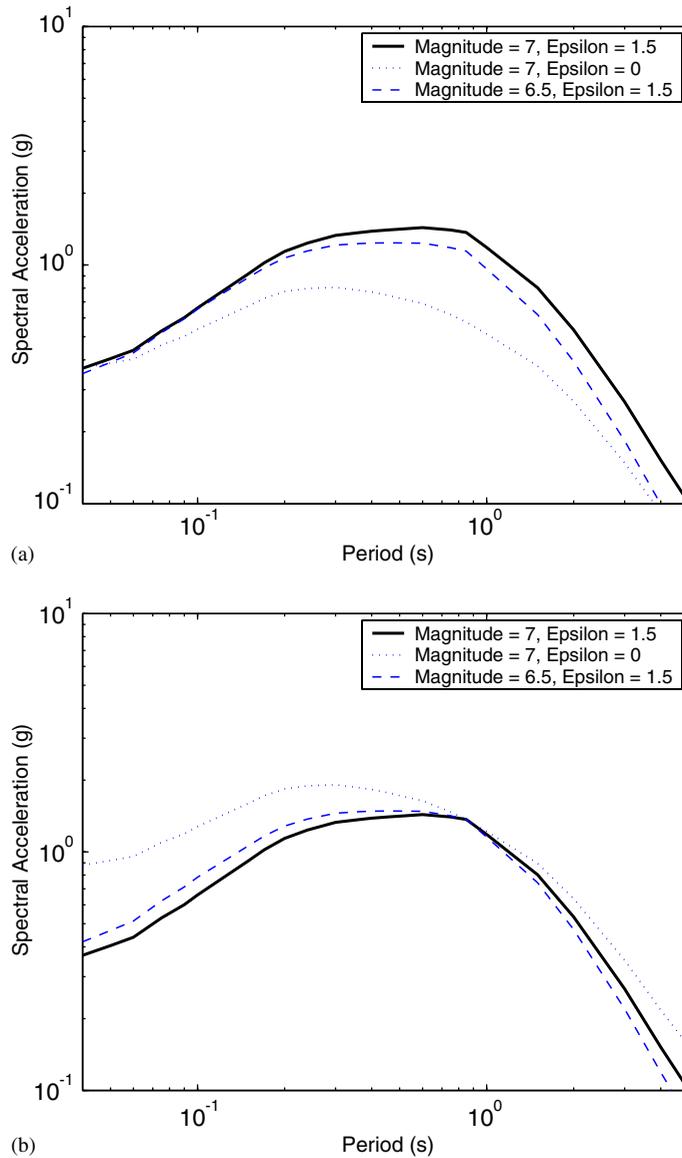


Figure 5. (a) Expected response spectra for three scenario events; and (b) expected response spectra for three scenario events, scaled to have the same $S_a(0.8\text{ s})$ value.

were also examined. However, neither of these measures showed significant improvement over ε , and thus were rejected in favor of the simpler ε parameter.

For these reasons, the ε predictor is believed to have advantages over alternative response predictors, making it an interesting candidate as an effective intensity measure.

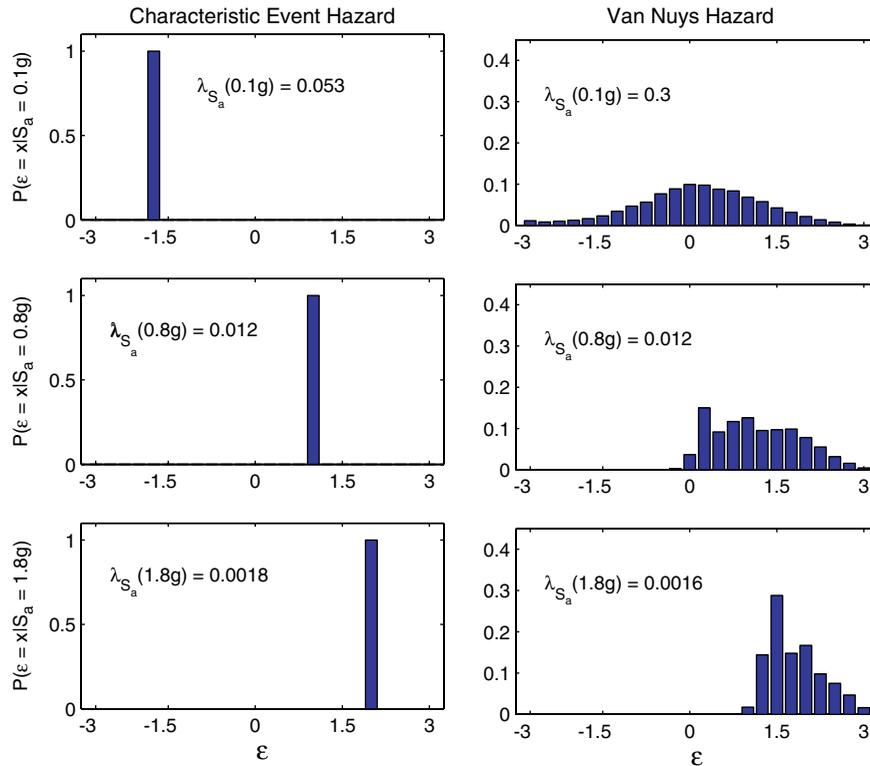


Figure 6. Disaggregation of PSHA results. The conditional distribution of ε given $S_a(0.8s) = x$ is shown for both fault models at three different hazard spectral acceleration levels associated with three different mean annual frequencies of exceedance.

Epsilon and ground motion hazard

Having established ε as an effective predictor of structural response supplemental to $S_a(T_1)$, it is useful to consider what values of ε are typically to be anticipated in ground motions that are of interest to structural engineers. For a given site and a given fault, the three parameters that can vary are magnitude, distance, and ε . In a Probabilistic Seismic Hazard Analysis, the possible values of these three parameters and their likelihoods are integrated over for each fault, and the hazard contributions of all faults are summed [29]. The result is a curve specifying the annual rate of exceeding varying levels of ground motion intensity. For purposes of illustration, consider a hypothetical site that has a single fault, capable of producing only magnitude 6.5 events at a distance of 8 km. At this site, magnitude and distance are fixed for all events, so ε is the only free variable in the PSHA analysis; thus the effect of ε can be more clearly seen. A single event model is also quite representative of the ground motion hazard situation at many sites located near a single large fault (i.e. some urban locations near the San Andreas or Hayward faults in northern California). This model will be referred to as the ‘Characteristic Event model’.

With this model, the median value of spectral acceleration (i.e. the exponential of the mean of $\ln S_a$) is $0.46g$, so any ground motions larger than $0.46g$ have a positive ε value. Thus, at low annual frequencies (where S_a values greater than $0.46g$ are seen) the ground motion hazard is governed exclusively by records with positive epsilons. This is seen in Figure 6: as the ground motion level is increased, the ε value seen in the disaggregation shows a corresponding increase.

Now consider the ground motion hazard at a real site surrounded by several faults, each of which is capable of producing events with a variety of magnitudes and distances. The hazard assessment used here is that for Van Nuys, California, at the site of the example structure which will be analyzed below. (The activity rate of the Characteristic Event model is specified such that both of these models have the same ground motion level at the 10% in 50 year hazard: $S_a(0.8\text{ s}) = 0.6g$. In addition, the magnitude and distance values for the Characteristic Event equal the mean values of the Van Nuys disaggregation at the 10% in 50 year level.) The hazard there is simply a summation of many hazard contributions each similar to the Characteristic Event model. There is a limit on the maximum magnitude and minimum distance of an event, so it is still true that large ground motion events will have positive ε values. In fact, as noted earlier, we can determine the ε values at a given hazard level by examining the disaggregation of the PSHA results. The disaggregation-based distribution of ε at the Van Nuys site is shown in Figure 6 for several levels of S_a . As the annual rate of exceedance decreases (i.e. as the ground motion level increases) the epsilons contributing to the hazard are seen to shift to larger values. Thus in general, for any hazard environment, the ground motion hazard at long low annual frequencies will be dominated by positive ε events.

EPSILON AND DRIFT HAZARD

In the previous section, it was established that at ground motion intensity levels with low annual rates of exceedance, records tend to have positive ε values. Further, we have seen that for a given S_a value, records with larger ε values cause systematically smaller responses in structures, due to the fact that ε is an indicator of peaks and valleys in the response spectra at the period of interest. But a typical random sample of records would have an average ε value of zero. Now consider the usual practice of using simply a scalar S_a with a suite of (on average) zero-epsilon records. When these records are used to estimate the response of a structure at (low annual frequency) high ground motion levels (typically characterized by positive epsilons), estimates of the frequency of exceeding large structural drifts are likely to be conservatively biased. This expectation can be confirmed by using the drift hazard procedure outlined earlier. The traditional approach using Equation (1) and $S_a(T_1)$ alone as the intensity measure is referred to here as the 'scalar-based approach'. The alternative approach using a vector-valued IM consisting of $S_a(T_1)$ and M or ε , and using Equation (7), is referred to as the 'vector-based approach'. A complete drift hazard curve is computed for a variety of structures using both of these approaches, and it is seen that neglecting to account for the ε value of a record nearly always results in over-estimation of the mean annual rates of exceeding large levels of drift.

Description of the structures analyzed

The primary structure analyzed is a reinforced-concrete moment-frame building. The building has 1960's era construction and is serving as a test-bed for PEER research activities [30]. The actual structure is located in Van Nuys, California, at the same site for which the ground motion hazard assessment above was conducted. A 2D model of the transverse frame created by Jalayer [31] is used here. This model has a first-mode period of 0.8 s, and contains non-linear elements that degrade in strength and stiffness, in both shear and bending [32]. Forty historical earthquake ground motions from California are used to analyze this structure. The events range in magnitude from 5.7 to 7.3, and the recordings are at distances between 6 and 36 km. Directivity effects are not expected at this site [30], so these effects were avoided by choosing records with small distances only when the rupture/site geometry suggested that near-fault effects would be unlikely; the ground motion velocity histories were not observed to contain pulse-like intervals. Epsilon was not calculated before the records were selected—therefore they have been effectively randomly selected with respect to ε . The set of 40 records was then scaled to 16 levels of $S_a(T_1)$ between $0.1g$ and $2.4g$.

To supplement the data from this structure, a series of generic frame structures was evaluated as well. Fifteen generic frame models were analyzed, with a variety of configurations, periods, and degradation properties. The specific model parameters are summarized in Table II. All of the structures are single-bay frames, with stiffnesses and strengths chosen to be representative of typical structures. Five structural configurations were considered, with varying numbers of stories and first-mode periods. A set of non-degrading models designed and analyzed by Medina and Krawinkler [33] was considered. These models do not have degrading elements, but the more flexible structures still have the potential to collapse due to $P-\Delta$ effects. A set of degrading models designed and analyzed by Ibarra [34] was also considered. These structures are identical to the models of Medina and Krawinkler, except for incorporation of elements that degrade in stiffness and strength. For each of the five building configurations, a non-degrading model and two degrading models were considered. A second set of forty records were used by those authors to analyze these structures, ranging in magnitude from 6.5 to 6.9, and ranging in distance from 13 to 40 km. It is difficult to generalize conclusions to all possible structures, but it is believed that by considering this wide range of models, the consistent effect of ε is apparent.

Drift hazard results

The first case considered is the reinforced concrete frame structure. This structure was evaluated for both the Characteristic Event and the Van Nuys ground motion hazard environments. The results are shown in Figure 7. We see that in both cases, versus the scalar approach, inclusion of ε in the intensity measure results in lower mean annual frequencies for high levels of drift. While inclusion of M with S_a has the same effect, it is much less pronounced, as anticipated.

This same procedure was repeated for the 15 generic frames considered. The drift hazard curves look similar to those of Figure 7. In order to present the results in a concise way, two important values for each structure are considered: the annual rate of exceeding 10% maximum interstorey drift ratio (this amount of drift could be interpreted to indicate collapse of the structure), and the drift that has a 10% chance of being exceeded in 50 years (i.e. the drift level z such that $\lambda(EDP \geq z) = 0.0021$). The percentage change in these values between

Table II. Percentage change in mean annual collapse rate and in the 10% in 50 year drift demand on a series of structural models when using the improved vector-based procedure versus the scalar-based procedure.

Number of stories	Period (s)	Degradation model	Reduction in mean annual rate of collapse	Reduction in drift at 10% in 50 year hazard
3	0.3	a	55%	93%
3	0.3	b	58%	91%
3	0.3	none	0%	17%
3	0.6	a	40%	<i>n.a.</i>
3	0.6	b	51%	73%
3	0.6	none	99%	25%
7	0.8	c	43%	<i>n.a.</i>
9	0.9	a	72%	42%
9	0.9	b	80%	41%
9	0.9	none	99%	32%
9	1.8	a	5%	18%
9	1.8	b	49%	20%
9	1.8	none	71%	25%
15	3.0	a	41%	5%
15	3.0	b	40%	<i>n.a.</i>
15	3.0	none	46%	13%

Degradation models:

- (a) Ibarra degradation parameters: peak oriented model,
 $\delta_c/\delta_y = 4$, $\alpha_c = -0.10$, $\alpha_s = 0.03$, $\gamma_{s,c,k,a} = \infty$
- (b) Ibarra degradation parameters: peak oriented model,
 $\delta_c/\delta_y = 4$, $\alpha_c = -0.05$, $\alpha_s = 0.03$, $\gamma_{s,c,k,a} = 50$
- (c) Pincheira model, with parameters calibrated to reflect concrete connections with representative detailing [19]

the scalar-based (S_a) result and the vector-based (S_a, ε) result is listed in Table II. For some cases, the structure collapses with a greater than 10% probability in a 50 year period (e.g. this is seen in Figure 7). In these cases, there is no drift level associated with the 10% in 50 year hazard level, so the corresponding cell in Table II is marked 'n.a.' (not applicable). In addition, the non-degrading structure with a period of 0.3 s is not predicted to collapse at the ground motion levels present at the site, so there is no change in the frequency of collapse. We see that in every case (besides these special cases) the improved vector-based procedure produces lower demands on the structure. One interesting insight: it can be noted that the collapse rate reduction is typically less for the longer period version of otherwise similar models; the absolute rates are, however larger in such cases implying smaller values of ε , and hence less effect.

DISCUSSION

It has been shown that ε has a systematic effect on the shape of a record's response spectra, and thereby an important effect on the response of non-linear MDOF models. This being the case, it is desirable to account for the effect of ε when predicting the annual rate of exceeding

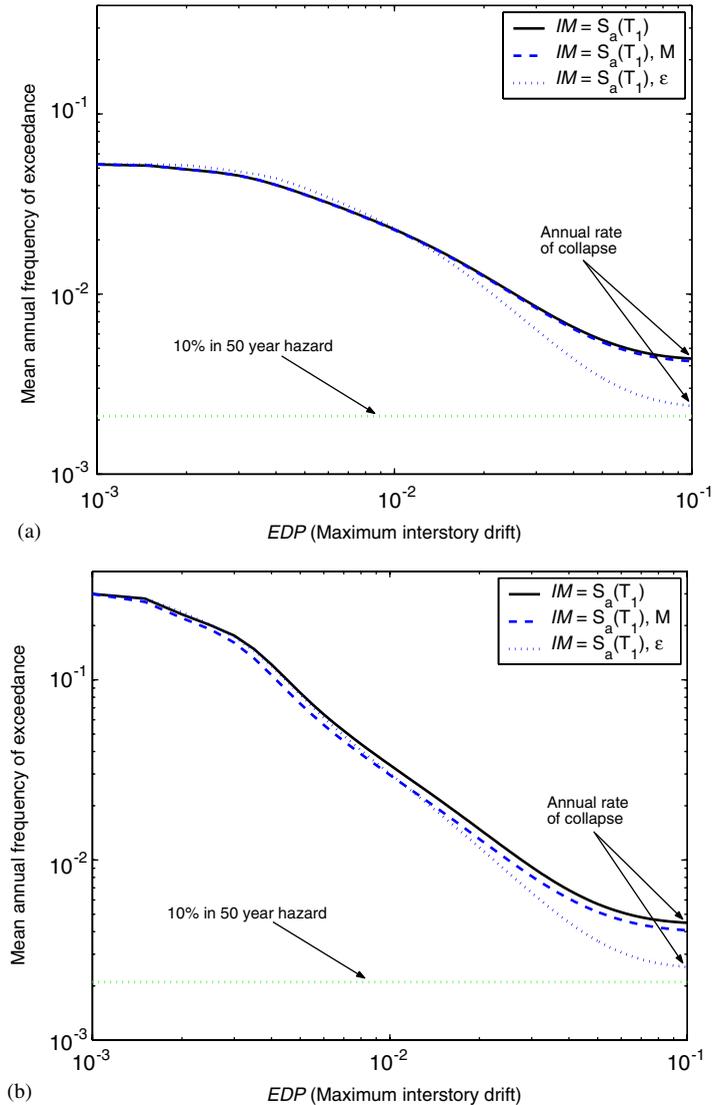


Figure 7. Seven storey reinforced concrete frame. Mean annual frequency of exceedance versus maximum interstorey drift: (a) scalar-based and vector-based drift hazard curves for the characteristic event hazard; and (b) scalar-based and vector-based drift hazard curves for the Van Nuys hazard.

a given drift level. There is more than one way to accomplish this. One approach is to consider ε values carefully when selecting ground motions to use in analysis. In current practice, ground motions are selected so that they match the magnitude and distance values of the events that dominate the disaggregation, and the soil conditions present at the site under consideration [10]. This paper suggests that ε should also be considered when selecting ground motions. Unfortunately, the existing library of recorded ground motions is not large enough such that

all desired parameters can be matched simultaneously, especially for a sample of nominal size. In light of the relatively greater effect that ε has on structural response, the desire to closely match distance and magnitude should probably be relaxed in favor of matching ε levels. For example, one could match ε while also trying to match magnitude, but allow records from a wide range of distances to be used. When selecting records to match ε , one needs to remain aware that target ε values will increase as the mean annual frequency considered decreases (the target magnitude and distance may also change as the annual frequency decreases, and this is sometimes accounted for in the selection process, e.g. Reference [10]).

The method to address the effect of ε proposed here is to adopt an *IM* that accounts for the effect of ε (i.e. an *IM* that is sufficient with respect to ε [35]). The most obvious *IM* that accomplishes this is the vector-valued *IM* presented in this paper, with $S_a(T_1)$ and ε as parameters. One first assesses the dependence of drift on ε at one or more levels of S_a , as described herein. One level at or near the mean annual frequency of interest (e.g. 2% in 50 years) may be sufficient in some cases. The sample size can probably be limited to the order of 10 records, especially if record selection is designed to capture, for example, the dependence of ε in the 0 to +2 range (note that with this approach, a wide range of ε values in the suite of records is desirable in order to improve the slope fit, in contrast to the previous approach where only records with specific ε values are desired). Then the drift hazard curve is computed using the vector-based procedure of Equation (7), as was done to produce Figure 7.

Finally, there is a possibility that alternative scalar intensity measures may account for the effects of ε more sufficiently than spectral acceleration (for instance, some of the improved scalar *IMs* that have been proposed recently [25, 35, 36]). If this were the case, then ε would not need to be included as a second parameter of the *IM*. However, this hypothesis remains to be tested. In the meantime, the most direct way to address the observed effect of ε is to use the vector-valued *IM* presented in this paper.

CONCLUSIONS

A method for calculating the probabilistic response of structures with a vector-valued *IM* is used to evaluate the significance of magnitude, distance and ε on the response of structures, conditioned on spectral acceleration. It is seen that ε has a significant effect on the response of structures, because it is an indicator of spectral shape (more specifically, it tends to indicate whether S_a at a specified period is in a peak or a valley of the spectrum). For a fixed $S_a(T_1)$, records with positive ε values cause systematically smaller demands in structures than records with negative ε values. The effect of ε on structural response given $S_a(T_1)$ is seen to be greater than the effect of magnitude or distance.

In addition, by examining disaggregation of the ground motion hazard, it is seen that at low mean annual frequency of exceedance the ground motions are all positive-epsilon motions. Therefore, the practice of scaling up zero-epsilon (on average) records to represent records with positive epsilons is likely to result in over-estimation of the demand on the structure. This will lead to over-estimation of drift at a given hazard level, or over-estimation of the mean annual frequency of collapse.

A vector-valued *IM* consisting of $S_a(T_1)$ and ε has been proposed in this paper. The proposed *IM* will account for the effect of ε on structural response. Alternatively, one could

correct for the effect of ε on structural response by intelligently choosing records that have the proper ε value, and then using $S_a(T_1)$ as a scalar IM .

The results presented here are based on a suite of structural models that are similar in behavior to many frame structures. However, these models are by no means representative of all classes of structures in existence. It is believed that the effect of ε will be seen in a broad class of structures, but further work is needed to further confirm and quantify this effect.

ACKNOWLEDGEMENTS

This work was supported primarily 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. We thank Dr Ricardo Medina and Dr Luis Ibarra for providing structural analysis results for the generic frame structures used. We also thank Dr Norman Abrahamson for allowing us to use his PSHA code to perform the ground motion hazard analysis.

REFERENCES

1. Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. *Earthquake Spectra* 1998; **14**(3):469–500.
2. Abrahamson NA. Statistical properties of peak ground accelerations recorded by the Smart 1 array. *Bulletin of the Seismological Society of America* 1988; **78**:26–41.
3. Abrahamson NA. State of the practice of seismic hazard evaluation. *GeoEng 2000*. Melbourne, Australia, 2000.
4. Abrahamson NA, Silva WJ. Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismological Research Letters* 1997; **68**:94–126.
5. Boore DM, Joyner WB, Fumal TE. Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work. *Seismological Research Letters* 1997; **68**:128–153.
6. Cornell CA, Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000; **3**. <http://peer.berkeley.edu/news/2000spring/performance.html>
7. Bazzurro P, Cornell CA, Shome N, Carballo JE. Three proposals for characterizing MDOF nonlinear seismic response. *Journal of Structural Engineering* (ASCE) 1998; **124**(11):1281–1289.
8. Jalayer F, Beck JL, Porter KA. Effects of ground motion uncertainty on predicting the response of an existing RC frame structure. *13th World Conference on Earthquake Engineering*. 2004. Vancouver, Canada, 10p.
9. Baker JW, Cornell CA. Choice of a vector of ground motion intensity measures for seismic demand hazard analysis. *13th World Conference on Earthquake Engineering*. 2004. Vancouver, Canada, 15p.
10. Stewart JP, Chiou S-J, Bray JD, Graves RW, Somerville PG, Abrahamson NA. Ground motion evaluation procedures for performance-based design. *PEER 2001-09*. 2001, Pacific Earthquake Engineering Research Center, University of California at Berkeley: Berkeley, California. 229p.
11. Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. *Applied Linear Statistical Models*; 4th edn. McGraw-Hill: Boston, 1996; 1408p.
12. Shome N. *Probabilistic Seismic Demand Analysis of Nonlinear Structures*. Department of Civil and Environmental Engineering, Stanford University, 1999; 320p. <http://www.stanford.edu/group/rms/>.
13. Aslani H, Miranda E. Probabilistic response assessment for building-specific loss estimation. *PEER 2003-03*. 2003, Pacific Earthquake Engineering Research Center, University of California at Berkeley: Berkeley, California. 49p.
14. Benjamin JR, Cornell CA. *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill: New York, 1970; 684p.
15. Bazzurro P, Cornell CA. Vector-valued probabilistic seismic hazard analysis. *7th U.S. National Conference on Earthquake Engineering*. 2002, Boston, MA: Earthquake Engineering Research Institute. 1 CD-ROM.
16. McGuiire RK. Probabilistic seismic hazard analysis and design earthquakes: Closing the loop. *Bulletin of the Seismological Society of America* 1995; **85**:1275–1284.
17. Bazzurro P, Cornell CA. Disaggregation of seismic hazard. *Bulletin of the Seismological Society of America* 1999; **89**:501–520.
18. Abrahamson NA. Personal communication, Pacific Gas & Electric Co., San Francisco, California, 2004.

19. U.S. Geological Survey Earthquake Hazard Maps. 2002. <http://eqhazmaps.usgs.gov/>.
20. Bazzurro P. *Probabilistic Seismic Demand Analysis*. Department of Civil and Environmental Engineering, Stanford University: Stanford, CA, 1998; 329p. <http://www.stanford.edu/group/rms/>.
21. Kennedy RP, Short SA, Merz KL, Tokarz FJ, Idriss IM, Power MS, Sadigh K. Engineering characterization of ground motion—Task I: Effects of characteristics of free-field motion on structural response. *NUREG/CR-3805*. 1984, U.S. Nuclear Regulatory Commission: Washington, D.C.
22. Chopra AK. *Dynamics of Structures: Theory and Applications to Earthquake Engineering*; 2nd edn. Prentice Hall: Upper Saddle River, NJ, 2001; 844p.
23. Iwan WD. Estimating inelastic response spectra from elastic response spectra. *Earthquake Engineering and Structural Dynamics* 1980; **8**:375–388.
24. Kennedy RP, Kincaid RH, Short SA. Prediction of inelastic response from elastic response spectra considering localized nonlinearities and soil-structure interaction. *8th SMIRT*. 1985, 427–434.
25. Cordova PP, Deierlein GG, Mehanny SSF, Cornell CA. Development of a two-parameter seismic intensity measure and probabilistic assessment procedure. *The Second U.S.–Japan Workshop on Performance-based Earthquake Engineering Methodology for Reinforced Concrete Building Structures*. 2001, Sapporo, Hokkaido. 187–206.
26. Vamvatsikos D. *Seismic Performance, Capacity and Reliability of Structures as Seen through Incremental Dynamic Analysis*. Department of Civil and Environmental Engineering, Stanford University: Stanford, CA, 2002; 152p. <http://www.stanford.edu/group/rms/>.
27. Inoue T, Cornell CA. Seismic hazard analysis of multi-degree-of-freedom structures. *RMS-8*. 1990, Reliability of Marine Structures: Stanford, CA. 70p.
28. Carballo JE. *Probabilistic Seismic Demand Analysis: Spectrum Matching and Design*. Department of Civil and Environmental Engineering, Stanford University: Stanford, CA, 2000; 259p. <http://www.stanford.edu/group/rms/>.
29. Kramer SL. *Geotechnical Earthquake Engineering*. Prentice-Hall Civil Engineering and Engineering Mechanics Series. Prentice Hall: Upper Saddle River, NJ, 1996; 653p.
30. Van Nuys Hotel Building Testbed Report: Exercising Seismic Performance Assessment. H. Krawinkler, Editor. 2004, Pacific Earthquake Engineering Research Center, University of California at Berkeley: Berkeley, California.
31. Jalayer F. *Direct Probabilistic Seismic Analysis: Implementing Non-Linear Dynamic Assessments*. Department of Civil and Environmental Engineering, Stanford University: Stanford, CA, 2003; 244p. <http://www.stanford.edu/group/rms/>.
32. Pincheira JA, Dotiwala FS, D'Souza JT. Seismic analysis of older reinforced concrete columns. *Earthquake Spectra* 1999; **15**:245–272.
33. Medina RA, Krawinkler H. *Seismic Demands for Nondeteriorating Frame Structures and Their Dependence on Ground Motions*. John A. Blume Earthquake Engineering Center Report No. 144. Stanford University: Stanford, CA, 2003; 347p.
34. Ibarra LF. *Global Collapse of Frame Structures under Seismic Excitations*. Department of Civil and Environmental Engineering, Stanford University: Stanford, CA, 2003; 324p.
35. Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra* 2005 (under revision of publication).
36. Mori Y, Yamanaka T, Luco N, Nakashima M, Cornell CA. Predictors of seismic demand of SMRF buildings considering post-elastic mode shape. *13th World Conference on Earthquake Engineering*. 2004. Vancouver, Canada, 15.