Information-Based Relative Sufficiency of Some Ground Motion Intensity Measures

F. Jalayer
University of Naples, Federico II, Italy

J. L. Beck
California Institute of Technology, USA

F. Zareian
University of California - Irvine, USA

SUMMARY:
The seismic risk assessment of a structure in performance-based design (PBD) may be significantly affected by the representation of ground motion uncertainty. In PBD, the uncertainty in the ground motion is often represented by a probabilistic description of a scalar parameter, or low-dimensional vector of parameters, known as the intensity measure (IM), rather than a full probabilistic description of the ground motion time history in terms of a stochastic model. In this work, a new procedure employing what we call a relative sufficiency measure is introduced based on information theory concepts in order to quantify the suitability of one IM relative to another in representing ground motion uncertainty. Based on this relative sufficiency measure, several alternative scalar- and vector-valued IMs are compared in terms of the expected difference in information they provide about a predicted structural response parameter, namely, the seismically-induced drift in an existing reinforced-concrete frame structure.

Keywords: Ground motion uncertainty, Intensity measure, Information theory, Relative entropy

1. INTRODUCTION

It is common to represent the uncertainty in the ground motion with a probabilistic model for a parameter, or a vector of a few parameters, related to the ground motion and known as an intensity measure (IM) (e.g. Luco and Cornell 2007; Jalayer and Cornell 2009; Goulet et al. 2007). One then faces the question of how suitable the adopted IM is for representing ground motion uncertainty.

Luco and Cornell (2007) have proposed sufficiency as one of the criteria for measuring the suitability of an IM in representing the dominant features of ground shaking. A sufficient IM has been defined as one that renders the structural response conditional on this IM to be independent of other ground motion characteristics. Establishing sufficiency in an absolute sense is likely to require high-dimensional vector IMs since it involves independence, conditional on the IM, of a designated structural response parameter from other ground motion characteristics for all possible values of the IM. Information theory concepts can be employed in order to measure the suitability of one IM relative to another in representing ground motion uncertainty. The (Shannon) entropy of an uncertain-valued variable is a measure of the amount of uncertainty in the value of that variable (Shannon 1948; Cover and Thomas 1991); more specifically, it is a measure of the missing information that is required (on average) to specify the value of the uncertain variable. In this work, based on the application of entropy and the related concept of relative entropy, we introduce a simple quantitative measure, called herein the relative sufficiency measure, for comparing the suitability of several IMs. This measure states (on average) how much more information about the designated structural response parameter one IM gives relative to another. We also present a case-study using the relative sufficiency measure to compare the suitability of various IMs for predicting the maximum inter-story drift ratio response in an existing reinforced-concrete moment-resisting frame located in Los Angeles.

2. METHODOLOGY

2.1 Information and Entropy

Let $P[A|I_A]$ denote the probability that statement A is true based on the information in statement I_A. The information measure $H[A|I_A]$ is a function of $P[A|I_A]$ that quantifies the amount of missing information about A given I_A and it is expressed [in bits] as:

$$H[A \mid I_A] = -\log_2 P[A \mid I_A] \tag{2.1}$$

Now let I_A state that one of $\{A_k: k=1,...,n\}$ is true and that $P[A_i \wedge A_j|I_A]=0$ if $i \neq j$, i.e., the A_k are mutually exclusive and exhaustive under I_A. Define the entropy (Shannon 1948) of the set of A_k as the mean of the missing information about $\{A_k: k=1, ..., n\}$, then:

$$H[\{A_k : k = 1,...,n\} \mid I_A] = \sum_{k=1}^{n} P[A_k \mid I_A] H[A_k \mid I_A] = -\sum_{k=1}^{n} P[A_k \mid I_A] \log_2 P[A_k \mid I_A] \tag{2.2}$$

where I_A specifies a probability model $P[A_k|I_A]$, k=1,...,n, for the A_k.

2.2 Relative Entropy

Suppose we gain additional (incomplete) information J_A about statement A, so that the probability $P[A_k|I_A]$ changes to $P[A_k|I_A \wedge J_A]$, then the *amount of information gain* about A from J_A over I_A is defined by:

$$\Delta H[A \mid I_A \mid J_A] = H[A \mid I_A] - H[A \mid I_A \wedge J_A] = \log_2 \frac{P[A \mid I_A \wedge J_A]}{P[A \mid I_A]} \tag{2.3}$$

which is the original missing information $H[A|I_A]$ minus the subsequent missing information $H[A|I_A,J_A]$. Note that ΔH may be negative, which signals an information loss. Define the *relative entropy* (Cover and Thomas 1991) of $\{A_k,\ k=1,\ ...,\ n\}$, n mutually exclusive and exhaustive statements, as the mean information gain from J_A over I_A:

$$D[\{A_k : k = 1,...,n\} \mid I_A \mid J_A] = \sum_{k=1}^{n} P[A_k \mid I_A \wedge J_A] \log_2 \frac{P[A_k \mid I_A \wedge J_A]}{P[A_k \mid I_A]} \tag{2.4}$$

It can be shown that the mean information gain when information J_A is added to information I_A is always non-negative; it is zero only when $P[A_k|I_A]=P[A_k|I_A,{\wedge}J_A]$, for all $k=1,\ ...,\ n$, i.e., information J_A does not change the probabilities $P[A_k|I]$ for all k (Cover and Thomas 1991). The relative entropy is also called the *divergence* or *Kullback-Leibler* information (Kullback 1959) or *cross entropy*.

2.3 Relative Entropy of a Continuous Variable

Let $A_k=\{X \in \Delta V(x_k)\}$ where $\Delta V(x_k)$ is a volume element at x_k of volume $|\Delta V(x_k)|$. Define the PDF for X:

$$p(x_k) = \lim_{\Delta V \to 0} \frac{P[A_k \mid I_A]}{|\Delta V(x_k)|} \tag{2.5}$$

The relative entropy can be calculated by partitioning the entire X-space with volume elements:

$$D(X \mid I_A \mid J_A) = \lim_{\Delta V \to 0} \sum_{k=1}^{n} q(x_k) \,|\Delta V(x_k)| \log_2 \frac{q(x_k)}{p(x_k)} = \int q(x) \log_2 \frac{q(x)}{p(x)} dx \tag{2.6}$$

where $q(x_k)$ is the PDF corresponding to $P[A_k|I_A \wedge J_A]$ defined as in Eq.(2.5).

2.4 Sufficiency of an Intensity Measure in an Absolute Sense

Luco and Cornell (2007) have introduced *sufficiency* as one of the criteria for assessing the suitability of an intensity measure. Here, a variation of their original definition is given: For a designated structural response parameter, such as the maximum inter-story drift θ_{max}, the intensity measure *IM* is *perfectly sufficient* if and only if:

$$p(\theta_{\max} | \underline{\ddot{x}}_g) = p(\theta_{\max} | IM(\underline{\ddot{x}}_g)) \tag{2.7}$$

for all ground motion (acceleration) time-histories, $\underline{\ddot{x}}_g$, that can happen at a site. Sufficiency in this absolute sense is an extremely strong condition for an intensity measure; it is unlikely that any scalar or low-dimensional vector *IM* satisfies this condition. Qualitatively speaking, it means that the *IM* provides as much information about θ_{max} as the entire ground motion time-history $\underline{\ddot{x}}_g$. If the structural modeling uncertainty is assumed to be negligible, a perfectly sufficient *IM* would fully determine the scalar structural response parameter θ_{max}.

2.5 Relative Entropy and Sufficiency of an Intensity Measure

The concept of relative entropy in Eq.(2.6) can be applied to measure the average information gained about the maximum inter-story drift θ_{max} when the available information about the ground motion is increased from knowing only the intensity measure *IM* to knowing the entire ground motion time-history $\underline{\ddot{x}}_g$:

$$\begin{aligned} D(\theta_{\max} | IM \mid \underline{\ddot{x}}_g) &= \int p(\theta_{\max} | IM(\underline{\ddot{x}}_g) \wedge \underline{\ddot{x}}_g) \log_2 \frac{p(\theta_{\max} | IM(\underline{\ddot{x}}_g) \wedge \underline{\ddot{x}}_g)}{p(\theta_{\max} | IM(\underline{\ddot{x}}_g))} d\theta_{\max} \\ &= \int p(\theta_{\max} | \underline{\ddot{x}}_g) \log_2 \frac{p(\theta_{\max} | \underline{\ddot{x}}_g)}{p(\theta_{\max} | IM(\underline{\ddot{x}}_g))} d\theta_{\max} \end{aligned} \tag{2.8}$$

where *IM* is irrelevant under conditioning on $\underline{\ddot{x}}_g$ and so the conditioning on *IM* can be dropped in that case. Since the relative entropy D($\theta_{max}|IM|\underline{\ddot{x}}_g$) is zero if and only if $p(\theta_{max}|\underline{\ddot{x}}_g)=p(\theta_{max}|IM(\underline{\ddot{x}}_g))$, the adopted *IM* is perfectly sufficient if and only if the relative entropy D($\theta_{max}|IM|\underline{\ddot{x}}_g$) is zero. It is noted that the relative entropy provides a quantified measure of the sufficiency of an *IM* in the sense that the farther away it is from zero, the less sufficient (less informative) that the *IM* is about θ_{max}.

2.6 Relative Sufficiency Measure Definition

The relative sufficiency of alternative *IM*s can be measured by comparing the difference between their corresponding relative entropies as given in Eq.(2.8). First, note that the difference between relative entropies corresponding to IM_1 and IM_2 may be expressed as:

$$D(\theta_{\max} | IM_1 \mid \underline{\ddot{x}}_g) - D(\theta_{\max} | IM_2 \mid \underline{\ddot{x}}_g) = \int p(\theta_{\max} | \underline{\ddot{x}}_g) \log_2 \frac{p(\theta_{\max} | IM_2(\underline{\ddot{x}}_g))}{p(\theta_{\max} | IM_1(\underline{\ddot{x}}_g))} d\theta_{\max} \tag{2.9}$$

Therefore, the difference between relative entropies is a functional of the ground-motion time history $\underline{\ddot{x}}_g$. Its expected value over all the ground motions that could happen at the site is defined here as the *relative sufficiency measure* for θ_{max} of IM_2 relative to IM_1:

$$
\begin{aligned}
I(\theta_{\max}| IM_2 \mid IM_1) &= E[D(\theta_{\max}| IM_1 \mid \underline{\ddot{x}}_g) - D(\theta_{\max}| IM_2 \mid \underline{\ddot{x}}_g)] \\
&= \int\int p(\theta_{\max}| \underline{\ddot{x}}_g)\log_2 \frac{p(\theta_{\max}| IM_2(\underline{\ddot{x}}_g))}{p(\theta_{\max}| IM_1(\underline{\ddot{x}}_g))} d\theta_{\max} \cdot p(\underline{\ddot{x}}_g)\cdot d\underline{\ddot{x}}_g
\end{aligned}
\tag{2.10}
$$

where $p(\underline{\ddot{x}}_g)$ is the PDF for the ground-motion time history at the site. If structural modelling uncertainty is ignored, for given $\underline{\ddot{x}}_g$, θ_{max} is known and is equal to $\theta_{max}(\underline{\ddot{x}}_g)$. This means that the probability density $p(\theta_{max} \mid \underline{\ddot{x}}_g)$ reduces to a Dirac delta function $\delta(\theta_{max}(\underline{\ddot{x}}_g))$.Therefore, after a few algebraic manipulations, the relative sufficiency measure $I(\theta_{max}|IM \mid \underline{\ddot{x}}_g)$ can be expressed as:

$$
\begin{aligned}
I(\theta_{\max}| IM_2 \mid IM_1) &= E[D(\theta_{\max}| IM_1 \mid \underline{\ddot{x}}_g) - D(\theta_{\max}| IM_2 \mid \underline{\ddot{x}}_g)] \\
&= \int \log_2 \frac{p(\theta_{\max}(\underline{\ddot{x}}_g) \mid IM_2(\underline{\ddot{x}}_g))}{p(\theta_{\max}(\underline{\ddot{x}}_g) \mid IM_1(\underline{\ddot{x}}_g))} \cdot p(\underline{\ddot{x}}_g)\cdot d\underline{\ddot{x}}_g
\end{aligned}
\tag{2.11}
$$

The relative sufficiency measure $I(\theta_{max}|IM_2| \ IM_1)$ can be interpreted as a measure of how much information on average is gained about the uncertain structural response parameter θ_{max} by knowing IM_2 instead of IM_1. If the logarithm is calculated in base two, the relative sufficiency measure is expressed in terms of *bits* of information. If the relative sufficiency measure is zero, this means that on average the two *IM*s provide the same amount of information about θ_{max}. In other words, they are *equally sufficient*. If the relative sufficiency measure is positive, this means that on average IM_2 provides more information than IM_1 about θ_{max}, so IM_2 is *more sufficient* than IM_1. Similarly, if the relative sufficiency measure is negative, IM_2 provides on average less information than IM_1 and so IM_2 is *less sufficient* than IM_1.

2.7 Calculation of the Relative Sufficiency Measure

In order to calculate the relative sufficiency measure using Eq.(2.11), both $p(\theta_{max}|IM)$ and $p(\underline{\ddot{x}}_g)$ are needed, so one has to choose probability models for the structural response for each candidate *IM* and for the ground motion time history. Strictly speaking, then, the relative sufficiency measure is conditional on these probability models, in addition to being conditional on the chosen structural model.

In this study, the probability model $p(\theta_{max}|IM)$ is selected by first choosing a set of real ground motion records. The structural response for each of these ground motion records is obtained by performing non-linear dynamic analyses. Taking $p(\theta_{max}|IM)$ as a lognormal probability density function, the two parameters (mean and standard deviation) of each distribution can be estimated using simple linear regression of structural response versus the corresponding *IM* (Luco and Cornell 1998, Jalayer and Cornell 2009). The procedure for calculating the relative sufficiency of a two-dimensional vector-valued *IM* denoted by $[IM_1,IM_2]$ with respect to a given reference *IM* is very similar to that of a scalar *IM*. The only difference is in the construction of the lognormal probability model $p(\theta_{max}|IM_1, IM_2)$ where multi-variate linear regression of the structural response versus $[IM_1,IM_2]$ is used in order to estimate the mean and standard deviation for $p(\theta_{max}|IM_1, IM_2)$. The second step in evaluating the relative sufficiency measure is to calculate the expectation in Eq.(2.11) over the possible ground motions at the site. Strictly, this requires a probability model $p(\underline{\ddot{x}}_g)$ for future ground motions at the structural site but a simple approximation is to replace the expectation by an average over a selected set of ground motion records. However, the resulting average may not be a good estimate of the expected value in Eq.(2.11) which strictly should take into account all the ground motions possible at the site, weighted by how likely each one is. It is shown later in the example results that this can be done using a stochastic ground motion model in conjunction with deaggregation of the seismic hazard at the site.

3. APPLICATION OF THE RELATIVE SUFFICIENCY MEASURE

The methodology described in the previous section is applied to an existing reinforced-concrete frame in order to compare the suitability of candidate intensity measures by calculating their relative sufficiency measures.

3.1. Model Structure: Longitudinal Frame of an Existing Building

The case-study building is a 7-story hotel in Van Nuys, California (34.221° N, 118.471° W) which has been studied by several researchers since the 1971 San Fernando Earthquake (e.g. Krawinkler 2005; Jennings 1971). The building footprint is 63 ft (3 bays) by 150 ft (8 bays), where the longitudinal direction is oriented east-west. The building is approximately 65 ft tall. The structural system is made of cast-in-place reinforced-concrete moment-resisting frames and flat plates. In the perimeter, the flat plate is combined with beams for additional lateral resistance. Due to old design of this building, columns have non-ductile detailing. In this study we use a two-dimensional model of the longitudinal direction of the Van Nuys building. In this direction, the building has four moment-resisting frames (i.e., two interior and two exterior moment-resisting frames), two of which with half of the mass of the building are modeled for computational efficiency (Figure 3.1). The model's first-mode period is T_1=1.5sec with an effective modal mass participation factor of 87%.

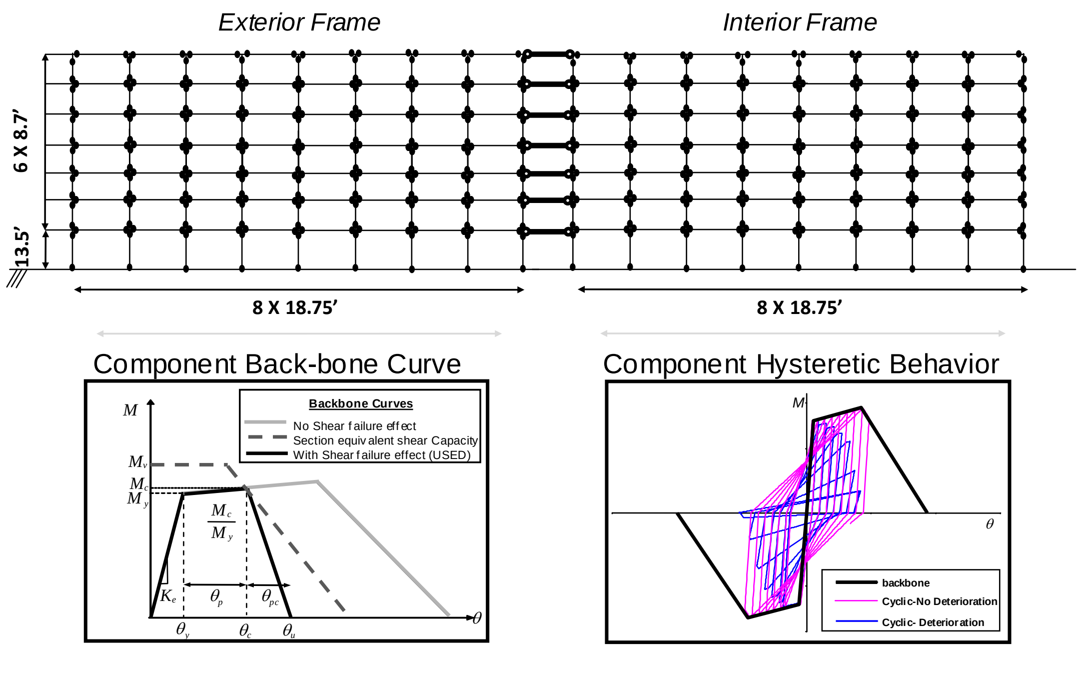

Figure 3.1. (Top) The structural model consisting of two moment-resisting frames in series. (Bottom-left) The component backbone curve; (Bottom-right) The component hysteretic behavior.

A critical damping coefficient of 5% in both the first and second longitudinal modes of vibration is considered for nonlinear response history analysis. The building model consists of structural component models that incorporate both monotonic and cyclic deterioration (Ibarra et al. 2005, see

Figure 3.1). Beams are modeled assuming that the effective width of the two-way slab is equal to the external frame's column strip. Column shear strengths were obtained from the recommendations made by Kowalsky and Priestley (2000), and assuming that the concrete contribution to shear strength is related to the flexural ductility demand in the hinge zone. In order to consider the shear transfer limitation in the interior flat slab to column joints, the maximum plastic rotation of the slab in the internal frame was limited to 0.04.

3.2. Selected Intensity Measures

Various alternative scalar intensity measures are compared in this study. One of the most commonly used *IM*s is *PGA* (peak of the ground motion acceleration time-history). Another widely-used *IM* is the spectral acceleration at the small-amplitude fundamental period T_1 of the structure, often denoted by $S_a(T_1)$, but more briefly referred to as the spectral acceleration S_a. Unlike *PGA*, which is only a characteristic of the ground motion, $S_a(T_1)$ also takes into account the ground-motion frequency content around the structure's first-mode period. Currently, this is the most widespread *IM* used in seismic risk analyses. Luco and Cornell (2007) have proposed a structure-specific intensity measure denoted by $IM_{1I,2E}$ that takes into account not only the ground-motion frequency content around the first two modal periods but also inelastic structural behavior to some extent:

$$IM_{1I,2E} = \frac{S_d^I(T_1, \xi_1, d_y)}{S_d(T_1, \xi_1)} \cdot IM_{12E} \tag{3.1}$$

and:

$$IM_{12E} = \sqrt{[PF_1 \cdot S_d(T_1, \xi_1)]^2 + [PF_2 \cdot S_d(T_2, \xi_2)]^2} \tag{3.2}$$

where PF_1 and PF_2 are modal participation factors for the first two modes of vibration; $S_d(T_1,\xi_1)$ and $S_d(T_2,\xi_2)$ are the spectral displacements with periods T_1 and T_2 and damping ratios ξ_1 and ξ_2 corresponding to the first two modes; and $S_d^I(T_1,\xi_1,d_y)$ is the spectral displacement of an elastic-perfectly plastic oscillator with period T_1 damping ratio ξ_1 and yield displacement d_y. Luco and Cornell (2007) have demonstrated that $IM_{1I,2E}$ is a better predictor of the structural response of moment-resisting frames than $S_a(T_1)$. The final scalar *IM* to be considered is S_a*, which is an *IM* proposed by Cordova et al. (2000) which takes into account spectral shape information:

$$S_a{*}(T_1, T_f) = S_a(T_1) \cdot \left(\frac{S_a(T_f)}{S_a(T_1)} \right)^{\alpha} \tag{3.3}$$

where T_f is another period at which spectral response is calculated. In this work, the values T_f and α are taken to be equal to T_2 and 0.5; i.e. $S_a{*} = [S_a(T_1)\, S_a(T_2)]^{0.5}$, a geometric mean of the spectral acceleration values at T_1 and T_2.

Three different vector-valued *IM*s are considered herein:

(a) [*PGA*,*M*] where *M* is the moment magnitude of the event generating the ground motion. It is expected that this vector is a better predictor than *PGA* alone;

(b) [$S_a(T_1)$, $S_a(T_f)$] where $T_f > T_1$ is another period at which spectral response is calculated;

(c) [$S_a(T_1)$, *epsilon*] where *epsilon* is the normalized regression residual for the ground motion prediction relationship that predicts spectral acceleration at the period T_1 in terms of the ground motion characteristics: magnitude M and source-to-site distance R_{rup}:

$$\ln S_a(T_1) = f(M, R_{rup}) + \sigma_{\ln S_a} \cdot epsilon \tag{3.4}$$

where f (M, R_{rup}) is the ground motion prediction relation and $\sigma_{\ln\ Sa}$ is the conditional logarithmic standard deviation for spectral acceleration at period T1 given the ground motion characteristics M and R_{rup}. Herein, *epsilon* is calculated based on f (M, R_{rup}) and $\sigma_{\ln\ Sa}$ from Abrahamson and Silva horizontal ground motion prediction relationship (Abrahamson and Silva 1997). Baker and Cornell (2005) have observed that *epsilon* is a good proxy for spectral shape because it can predict whether for a given period T_1, $S_a(T_1)$ is in a peak or valley of the spectrum. This means that, given $S_a(T_1)$, ground motion records with a positive *epsilon* (peak) lead to smaller demands compared to ground motions with a negative *epsilon* (valley). Therefore, the vector consisting of the pair of *IM*s [$S_a(T_1)$, *epsilon*] is expected to be a better *IM* for predicting the structural response than $S_a(T_1)$ alone.

3.3 Selected Ground Motion Records

The non-linear dynamic analyses are performed on a suite of 30 real ground-motion records that are selected from a ground motion database (PEER NGA 2010). The records are on stiff soil from a magnitude range of $6.5 \le M \le 7$ and source-to-site distances of $15 \le R \le 35$ km. For each ground motion, the structural response θ_{max}, as well as the values of the 8 candidate *IM*s are calculated.

3.4 Parameters for Probability Model p(θmax| IM)

A non-linear dynamic procedure referred to as the *Cloud Method* by Jalayer and Cornell (2009) has been employed in order to calculate the parameters of the lognormal PDF, $p(\theta_{max}|IM)$. The cloud method consists of first applying a suite of ground motion records to the structure and calculating the structural response θ_{max}. The parameters for the lognormal distribution can then be estimated by performing a simple linear regression on $\ln\theta_{max}$ versus the candidate *IM*. More specifically, the expected value of $\ln\theta_{max}$ given *IM* is modeled by a regression equation of the following form:

(a) For a scalar *IM*: $E[\ln\theta_{max}|IM] = a + b\ln(IM)$

(b) For $IM=[S_a(T_1),S_a(T_2)]$: $E[\ln\theta_{max}|IM] = a + b\ln S_a(T_1) + c\ln S_a(T_2)$

(c) For $IM=[PGA, M]$: $E[\ln\theta_{max}|IM] = a + b\ln PGA + cM$

(d) For $IM=[S_a(T_1), epsilon]$: $E[\ln\theta_{max}|IM] = a + b\ln S_a(T_1) + c\ epsilon$

In all cases, the standard deviation of $\ln\theta_{max}$ given *IM* is estimated by the standard error *s* of the regression. The estimated parameters *a*, *b*, *c* (when applicable) and *s* for each *IM* are given in Table 3.1.

3.5 Calculation of the Relative Sufficiency Measure for Selected Intensity Measures

In this section, the relative sufficiency measure is calculated in an approximate manner based on the set of recorded ground motions. As explained in the previous section, first the set of recorded ground motions is used to construct the probability models for the structural response θ_{max} given each candidate intensity measure. In the next step, the relative sufficiency measure in Eq. (2.11) is calculated in an approximate manner as the average of the logarithmic term inside the integral over the set of recorded ground motions.

The reference *IM* is taken to be $S_a(T_1)$ and the relative sufficiency measure for the other seven *IM*s relative to $S_a(T_1)$ is first estimated by simply replacing the expectation (an integral over all possible ground motion records) in Eq.(2.11) by an average over the set of selected ground motions. The relative sufficiency measures estimated in this way are listed in the second column of Table 3.2. The results can be interpreted for the studied building as, for example, *PGA* gives (on average) 1.33 bits of information *less* about the structural response θ_{max} than $S_a(T_1)$ while IM_{12E} gives (on average) 0.12 bits of information *more* about θ_{max} than $S_a(T_1)$. This ranks $IM_{1I,2E}$ as the least sufficient and [$S_a(T_1)$, $S_a(T_2)$] as the most sufficient of the *IM*s.

Table 3.1. Regression parameters for the adopted *IM*s. The regression related to [*PGA*, *M*] is performed on *M* and not on ln(*M*); the same is true for the regression on *epsilon*.

IM	a	b	c	s
$S_a(T_1)$	0.0355	0.82	NA	0.1444
PGA	0.0277	0.8309	NA	0.3656
Sa*(T1,T2)	0.0254	0.85	NA	0.1824
$IM_{1I,2E}$	0.0487	0.5393	NA	0.3945
IM_{12E}	0.1695	0.837	NA	0.1334
[$S_a(T_1)$, $S_a(T_2)$]	0.0334	0.7062	0.1433	0.1333
[PGA, M]	3.58E-04	0.7319	0.625	0.3576
[$S_a(T_1)$, epsilon]	0.0423	0.9048	-0.0615	0.1457

Although the approximation of Eq.(2.11) by an average over a set of ground motion records is straightforward to calculate, these estimates of the relative sufficiency measure may be too crude because the suite of ground motion records selected is not a random sample drawn from an appropriate PDF for the ground motion at the site. This deficiency in the approximation is confirmed by a more refined calculation of the relative sufficiency measure in Eq.(2.11), which is performed next.

3.6 Refined Calculation of Relative Sufficiency Measure

The expectation in the definition of the relative sufficiency measure should be calculated over the range of all possible ground motions at the site. This can be achieved by expanding the right-hand side of Eq.(2.11) with respect to source-to-site distance *R* and moment-magnitude *M* using the total probability theorem (Benjamin and Cornell 1970):

$$I(\theta_{\max} \mid IM_2 \mid IM_1) = \int \log \frac{p(\theta_{\max}(\underline{\ddot{x}}_g) \mid IM_2(\underline{\ddot{x}}_g))}{p(\theta_{\max}(\underline{\ddot{x}}_g) \mid IM_1(\underline{\ddot{x}}_g))} \cdot p(\underline{\ddot{x}}_g \mid M, R) \cdot p(M, R) dM dR d\underline{\ddot{x}}_g \tag{3.5}$$

The integration in Eq.(3.5) can be carried out using a standard Monte Carlo simulation scheme. This paper employs the deaggregation of seismic hazard (McGuire 1995; Bazzurro and Cornell 1998) at different levels of ground motion intensity in order to obtain a joint probability distribution $p(M,R)$ for magnitude and distance (Jalayer and Beck 2008). The stochastic ground motion model proposed by Atkinson and Silva (2000) is used to obtain the PDF $p(\underline{\ddot{x}}_g|M,R)$ for the ground motion time history given *M* and *R*. The simulation has been carried out using 2000 sample analyses and the resulting values for the relative sufficiency measure are presented in the third column of Table 3.2. Upon screening the maximum inter-story drift ratios calculated for the 2000 synthetic records, 6 cases of collapse or numerical non-convergence are detected. The relative sufficiency measures are calculated without considering these cases, and so are conditional on structural collapse not taking place.

Table 3.2. Relative sufficiency measures for alternative *IM*s relative to $S_a(T_1)$

Relative Sufficiency Measure	**Approximate**	**Refined**
$I(\theta_{max}\|PGA\|S_a(T_1))$	-1.330	-1.020
$I(\theta_{max}\|S_a{*}(T_1,T_2)\|S_a(T_1))$	-0.338	**0.542**
$I(\theta_{max}\|IM_{1I,2E}\|S_a(T_1))$	-1.450	-1.620
$I(\theta_{max}\|IM_{12E}\|S_a(T_1))$	**0.115**	**0.560**
$I(\theta_{max}\|[S_a(T_1), S_a(T_2)]\|S_a(T_1))$	**0.140**	**0.330**
$I(\theta_{max}\|[PGA, M]\|S_a(T_1))$	-1.280	-0.015
$I(\theta_{max}\|[S_a(T_1), \textit{epsilon}]\|S_a(T_1))$	**0.011**	**NA**

The results rank $IM_{1I,2E}$ as the least sufficient and rank IM_{12E} as the most sufficient followed closely by $S_a{*}(T_1,T_2)$, which is slightly different from the conclusion drawn before by taking the average over the suite of ground motion records. However, the results here are more defensible than those calculated previously using the simple average over the set of recorded ground motions. As can be seen from the results, the vector [PGA,M] gives almost the same amount of information about θ_{max} as $S_a(T_1)$; this is different from the conclusion drawn based on the approximate solution using the set of real ground motion records. However, in that case the residuals of the θ_{max}-PGA regression show very little or no trend with respect to the moment magnitude. This may be why the approximate calculation ranks [PGA,M] so poorly but the refined calculation ranks [PGA,M] almost as good as $S_a(T_1)$ for predicting θ_{max}. Looking at Table 3.2, it can be observed that the relative sufficiency measure corresponding to [$S_a(T_1)$, *epsilon*] is not reported. This is because the ground motion records used in the refined calculations are not real recordings.

3.7 Discussion

It may seem surprising that $IM_{1I,2E}$, the *IM* proposed by Luco and Cornell (2007), is ranked so poorly by both the approximate and refined method for the calculation of the relative sufficiency measure. However, this case-study structure reaches the ultimate capacity at low drift ratios (i.e., 1.8%) and experiences a steep post-capping stiffness. This means that $IM_{1I,2E}$ which is calculated based on elastic perfectly plastic non-linear behavior, may not capture well the non-linear behavior in the structure. Therefore, in terms of dynamic response given that collapse does not take place, the structure is going to behave more-or-less linearly. This also explains why the reference *IM*, $S_a(T_1)$, does well. The best *IM*s, however, are those that not only take into account information related to the first mode of vibration but also the information related to the second mode of vibration of the structure: [$S_a(T_1)$, $S_a(T_2)$], IM_{12E} and $S_a{*}(T_1,T_2)$.

4. CONCLUSIONS

A measure of the relative sufficiency of alternative intensity measures for representing ground motion uncertainty is derived in this work based on information theory concepts. This *relative sufficiency measure* quantifies the amount of information gained (on average) about a designated structural response parameter by adopting one intensity measure instead of another. Adopting the first-mode spectral acceleration $S_a(T_1)$ as the reference *IM*, the relative suitability of four scalar and three vector-valued *IM*s is quantified in terms of the relative sufficiency measure for the maximum inter-story drift ratio for a case-study building. It is found that the most sufficient (most informative) intensity measures are IM_{12E} and $S_a{*}(T_1,T_2)$, which give an average of 0.5 bit of more information than $S_a(T_1)$. The vector *IM* of *PGA* and magnitude *M* is just as sufficient (equally informative) as $S_a(T_1)$.

ACKNOWLEDGMENTS

This work was supported in part 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). 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 National Science Foundation. The first author also acknowledges the support from a George W. Housner Post-doctoral Fellowship in Civil Engineering from the California Institute of Technology.

REFERENCES

Abrahamson, N. A., Silva, W. J. (1997). “Empirical Response Spectral Attenuation Relations for Shallow Crustal Earthquakes.” Seismological Research Letters, 68(1), 94-127.

Atkinson, G. M., and Silva, W. (2000). “Stochastic modeling of California ground motions.” *Bulletin of the Seismic Society of America*, 90(2), 255-274.

Baker J. W., Cornell C. A. (2005). “A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon.” *Earthquake Engineering and Structural Dynamics*, 34:1193–217.

Benjamin, J. R., and Cornell, C. A. (1970). “Probability, Statistics and Decision for Civil Engineers.”, McGraw-Hill, New York, NY.

Bazzurro, P. and Cornell, C. A. (1998). “Disaggregation of seismic hazard.” *Bulletin of the Seismological Society of America*, 89(2), 501-520.

Cordova, P. P., Deierlein, G. G., Mehanny, S. S. F., and Cornell C. A. (2000). “Development of a two-parameter seismic intensity measure and probabilistic assessment procedure.” PEER-2000/10: *The Second U.S.-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures*, 11-13 September 2000, Sapporo, Hokkaido, Japan, Pacific Earthquake Engineering Research Center, University of California, pp. 195-214.

Cover, T. M., and Thomas, J. A. (1991). “Elements of Information Theory.”, John Wiley & Sons, Inc., New York, NY.

Goulet, C. A., Haselton, C. B., Mitrani-Reiser, J., Beck, J. L., Deierlein, G. G., Porter, K. A., Stewart, J. P. (2007). “Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building—from seismic hazard to collapse safety and economic losses.” Earthquake Engineering and Structural Dynamics, 36(13), 1973-1997.

Ibarra, L. F, Medina, R. A, and Krawinkler, H. (2005). “Hysteretic models that incorporate strength and stiffness deterioration.” *Earthquake Engineering and Structural Dynamics*, 34(12), 1489-1511.

Jennings, P. C. (1971). “Engineering Features of the San Fernando Earthquake of February 9, 1971.” Report EERL 71–02, California Institute of Technology, Pasadena, CA.

Jalayer, F., and Beck, J. L. (2008). “Effect of the alternative representations of ground motion uncertainty on seismic risk assessment of structures.” *Earthquake Engineering and Structural Dynamics*, 37(1): 61-79.

Jalayer F., and Cornell C. A. (2009). “Alternative nonlinear demand estimation methods for probability-based seismic assessments.” *Earthquake Engineering and Structural Dynamics*, 38(8): 951-972, 2009.

Kowalsky, M. J., Priestley, M. J. N. (2000). “Improved analytical model for shear strength of circular reinforced concrete columns in seismic regions.” *ACI Structural Journal*, 97(3): 388-396.

Krawinkler, H., ed. (2005). “Van Nuys Hotel building testbed report: exercising seismic performance assessment.” Report No. PEER 2005/11, Pacific Earthquake Engineering Research Center, University of California at Berkeley, Berkeley, California.

Kullback, S. (1959). “Information and Statistics.” J. Wiley, New York, NY.

Luco, N., and Cornell, C. A. (2007), “Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions.”, *Earthquake Spectra*, 23(2), 357-392.

Luco, N., and Cornell, C. A (1998). “Seismic drift demands for two SMRF structures with brittle connections.” *Structural Engineering World Wide 1998*, Elsevier Science Ltd., Oxford, UK, T158-3.

McGuire, R. K. (1995). “Probabilistic seismic hazard analysis and design earthquakes: closing the loop.”, *Bulletin of Seismic Society of America*, 85 (5), 1275-1284.

PEER NGA strong motion database, <http://peer.berkeley.edu/nga/> (25 November 2010).

Shannon, C. E. (1948). “A Mathematical Theory of Communication.”, Bell Syst. Tech. J., 27, 379-423, 623-656.