ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

Appraisal and mathematical properties of fragility analysis methods

S. Yi[a], K. G. Papakonstantinou[b], C. P. Andriotis[c], and J. Song[d]

[a] University of California, Berkeley, Berkeley, CA, USA, E-mail: visangri@berkeley.edu
[b] The Pennsylvania State University, University Park, PA, USA, E-mail: kpapakon@psu.edu
[c] Delft University of Technology, Delft, The Netherlands, E-mail: c.andriotis@tudelft.nl
[d] Seoul National University, Seoul, S. Korea, E-mail: junhosong@snu.ac.kr

ABSTRACT: Fragility analysis aims to compute the probabilities of a system exceeding certain damage conditions given different levels of hazard intensity. Fragility analysis is therefore a key process of performance-based earthquake engineering, with a number of approaches developed and widely recognized, including Incremental Dynamic Analysis (IDA), Multiple Stripe Analysis (MSA), and cloud analysis. Additionally, extended fragility analysis has recently been shown to possess important attributes of mathematical consistency and extensibility. This work provides a critical review of the different fragility methods by explaining the underlying probabilistic models and assumptions, as well as their connections to the extended fragility method. It is proven that IDA-based fragility curves provide an upper bound of the actual fragility, and cloud analysis manifests suboptimality issues arising from its underlying assumptions. MSA is identified to be a probit-linked Bernoulli regression model, similar to the one proposed by Shinozuka and coworkers. The latter, in turn, is shown to be a limiting subcase of the generalized linear model framework introduced within the extended fragility analysis. The paper first presents a simple case of one intensity measure and two damage condition states, and the discussion is subsequently extended to more general cases of multiple intensity measures and damage states. The discussed attributes are demonstrated in several numerical applications. Overall, this work aims to provide new insights on fragility methods, enabling efficient, accurate, and consistent estimations of structural performance, as well as promoting new research directions in earthquake engineering and other related fields.

KEYWORDS: Fragility analysis, extended fragility analysis, incremental dynamic analysis, multiple stripe analysis, cloud analysis, generalized linear models

1 INTRODUCTION

Fragility analysis aims to assess the probability of a system exceeding various damage states (DS) for a range of intensity measures (IM). It therefore has a central position in performance-based evaluation of structural systems subjected to hazards. A number of systematic fragility analysis methods have been proposed over the years (Baker 2015, Bakalis & Vamvatsikos 2018), including the recently developed extended fragility framework (Andriotis & Papakonstantinou, 2018), which can consistently handle multivariate IM and multiple DS of any dimensions. In Andriotis & Papakonstantinou (2018), it is proven that under broad probabilistic assumptions the softmax function describes a consistent mathematical form for fragility functions, and these functions should not be generally described by cumulative distribution functions over the intensity measures space, as many approaches currently suggest in the literature. In addition, extended fragility functions completely resolve, under no assumptions and limitations, the issue of crossings of fragility functions, which is a major mathematical inconsistency. In this paper, a critical review of available methodologies for fragility analysis is presented, and their theoretical attributes and connections to the class of extended fragility functions are explained. A unified viewpoint is therefore provided, based on derived

analytical results and formal proofs. Particularly, it is proven that Incremental Dynamic Analysis (IDA) (Vamvatsikos & Cornell 2002, 2004) represents a theoretical upper bound of the actual structural fragility, and that the cloud method only provides suboptimal fragility curves due to its inherent assumptions related to the utilized objective function. The Multiple Stripe Analysis (MSA) is studied as a subcase of the probit-linked Bernoulli model introduced by Shinozuka *et al*. (2000), which in turn is shown to be a limited subcategory of the extended fragility analysis framework. All relevant aspects, assumptions, eligibility, and relationships between methods are also provided for various settings, starting from the simplest case of one IM and two DS to general cases with multidimensional IM and multiple DS. In particular, the crossing issue between fragility curves of different threshold levels is discussed, and the fragility (hyper-)surface function for multidimensional IM is rigorously formulated by introducing generalized linear model theory.

2 FRAGILITY ANALYSIS

2.1 *Review on fragility analysis*

Fragility analysis is defined as the probability of system damage given the intensity of a hazard. Consider a single intensity measure parameter *IM*, and a binary damage state variable *DS*. The fragility function is represented as:

$$p_f(im) = \mathrm{P}(DS = 1 \mid IM = im) \tag{1}$$

where im denotes IM of interest, and DS takes values 1 or 0, depending on whether the damage is observed or not. Often, the DS is represented in terms of a critical response metric of the system, called the engineering demand parameter (EDP). In such cases, Eq. (1) can be rewritten as:

$$p_f(im) = \mathrm{P}(EDP \geq edp_o \mid IM = im) \tag{2}$$

where edp_o is a performance threshold value related to the analyzed system. Throughout the paper, we will consider a deterministic edp_o, but it is noted that the findings can be equally extended to cases with random thresholds.

2.2 *Lognormal-linked Bernoulli (Shinozuka et al.)*

A number of fragility methods including Shinozuka *et al*. (2000, 2003) have modeled the event of a system damage as a Bernoulli trial, represented as:

$$p(DS = z \mid im) = p_f(im)^{II\{z=1\}}(1 - p_f(im))^{II\{z=0\}} \tag{3}$$

where z is a binary variable that represents DS, and $II\{\bullet\}$ is an indicator function. The probability of system damage is defined by a link function $p_f(im)$, or the fragility function as the definition suggests. Particularly, Shinozuka *et al.* introduced a lognormal cumulative density function (CDF) as the link function:

$$p_f(im) = \Phi\left(\frac{\ln im - \theta}{\beta}\right) \tag{4}$$

Given data pairs of $\{im, z\}$, the parameters of Eq. (4) can be determined by maximum likelihood estimation (MLE) using the following likelihood function:

$$L = \prod_{n=1}^{N} p_{fn}^{\ II\{z_n=1\}}(1 - p_{fn})^{II\{z_n=0\}} \tag{5}$$

where $p_{fn}=p_f(im_n)$ and N is the number of data points. Hereafter, this method is referred to as Shinozuka *et al.*

2.3 *Multiple stripe analysis (MSA)*

MSA is performed at multiple discrete levels of intensity referred to as *stripes* (Jalayer and Cornell, 2009). Particularly, the MSA data are designed to form horizontal stripes on the $\{im, EDP\}$ space, and the dataset in each stripe represents different input hazard events under the same IM level. Considering only the samples in a stripe $IM=im$, the probability of observing k damaged cases out of total $n_s^{(im)}$ samples is regarded to follow a binomial distribution:

$$p(k \mid im) = \binom{n_s^{(im)}}{k} p_f(im)^k (1 - p_f(im))^{n_s^{(im)} - k} \tag{6}$$

For the relevant link function, MSA introduces again a lognormal CDF, as in Eq. (4). Similarly to the Shinozuka *et al.* method, the likelihood function is written as:

$$L = \prod_{j=1}^{N_{IM}} \binom{n_j}{k_j} p_{fj}^{\ z_j}(1 - p_{fj})^{n_j - k_j} \tag{7}$$

where $p_{fj}=p_f(im_j)$, $n_j=n_s^{(im_j)}$ and N_{IM} is the number of stripes. Since binomial and Bernoulli trials are mathematically equivalent and both methods are using a lognormal CDF as the link function, MSA

(Eq. 7) and Shinozuka *et al*. (Eq. 4) give identical results, when both are applied at the stripe-based dataset. Thus, MSA can be considered a subcategory of Shinozuka *et al*., which is more general in terms of the flexibility the utilized data can have. An advantage of MSA, is that it can also estimate the fragility non-parametrically at each stripe, as a ratio of damage-related and total samples. However, this requires availability of ample data per stripe.

2.4 *Cloud analysis*

Cloud analysis in Jalayer *et al.* (2018), Miano *et al.* (2017) and other works employs a power-law relation between IM and EDP to estimate the fragility. Particularly, cloud analysis first divides the simulation results to the collapsed (*C*) (for which EDP is considered infinite) and non-collapsed cases (*NC*) by introducing the logistic classification:

$$\mathrm{P}(C \mid im) = \frac{1}{1+e^{\beta_0+\beta_1 \ln im}} \tag{8a}$$

$$\mathrm{P}(NC \mid im) = 1-\mathrm{P}(C \mid im) \tag{8b}$$

and, then, a linear regression is performed in the log-space of non-collapsed samples:

$$E[\ln EDP \mid im, NC] = \alpha_0 + \alpha_1 \ln im \tag{9}$$

Under the power-law assumption, the probability exceeding the limit state, edp_o,is estimated as:

$$\mathrm{P}(EDP \geq edp_o \mid im, NC) = \Phi\left(\frac{edp_o - E[\ln EDP \mid im]}{\alpha_{res}}\right) \tag{10}$$

where α_{res}^2 is the sum of squared residuals. All the parameters in Eqs. (9) and (10) can be identified from a simple linear regression model with closed form expressions, whereas those of Eq. (8) require a separate MLE optimization that does not have analytical solution and is performed numerically. The two results are then combined to form a single fragility function, as:

$$p_f(im) = \mathrm{P}(EDP \geq edp_o \mid im, NC)\mathrm{P}(NC \mid im) + \mathrm{P}(C \mid im) \tag{11}$$

following the total probability theorem. Given the underlying and inherent limiting assumptions of Eqs. (9) and (10), the solution of the cloud analysis introduces several sources of model bias, eventually compromising accuracy.

2.5 *Incremental dynamic analysis (IDA)*

IDA transforms the main question of fragility analysis from "will the system fail given an intensity level of hazard? (Eq. (2))" into "what is the intensity level distribution at the onset of system failure?", i.e.,:

$$p_f(im) = \mathrm{P}(IM \leq im \mid EDP = edp_o) \tag{12}$$

Specifically, by identifying the probability distribution of the critical IM, which produces responses that meet the capacity of the system, the IDA reduces the statistical task to a simple 1-dimension probability distribution fitting problem (Vamvatsikos & Cornell 2002, 2004).

In general, acquiring the critical (or design) IM is an inverse problem that requires numerical optimization. To avoid such iterations, IDA introduces the so-called IDA curve, which represents a trajectory of structural responses for a sequence of IM scaling factors and estimates the IM value at the response limit by interpolation. For each ground motion, one IDA curve is generated in the space of IM and EDP, by scaling the ground motion intensity, and the point where the curve meets the threshold EDP level, edp_0, is considered as the critical IM. The collected samples of critical IM are used to identify the corresponding empirical CDF leading to the definition of fragility curve in Eq. (12). A lognormal CDF is also often chosen as a parametric probability distribution to fit Eq. (12):

$$p_f(im) = F_{IM|edp_o}(im \mid edp_o) = \Phi\left(\frac{\ln im - \theta_o}{\beta_o}\right) \tag{13}$$

One advantage of IDA is that it does not require numerical iterations for this parametric estimation in Eq. (13), since the parameters are simply the sample mean and standard deviation of the critical IM in the log-transformed space.

However, IDA curves can often exhibit negative slopes, leading to multiple threshold crossing points. When an IDA curve presents multiple threshold crossing points, the failure IM is defined as the lower one. This assumption leads to an overestimation of the fragility. In more detail, consider an IDA curve given by event α denoted as $g_\alpha(im)$. The IDA estimates are unbiased only if the conditions below are satisfied:

$$g_\alpha(im) \geq edp_o \quad \text{for all} \quad im \in [im_{\alpha,o}, \infty) \tag{14}$$

where $im_{\alpha,o}$ is the smallest IM on the IDA curve that crosses the threshold level (see Fig.1). In case Eq. (14) is not satisfied, the fragility curve obtained by Eq. (12) always overestimates the exact fragility definition in Eq. (2):

$$\begin{aligned}
&\mathrm{P}\left(EDP \ge edp_o \mid IM = im\right) \\
&= \mathbb{E}_A\left[\ \Pr\left(g_A(IM) \ge edp_o \mid IM = im, A = \alpha\right)\right] \\
&= \mathbb{E}_A\left[\ \Pr\left(g_\alpha(im) \ge edp_o\right)\right] \\
&= \mathbb{E}_A\left[\ \mathbb{I}\left(g_\alpha(im) \ge edp_o\right)\right] \\
&\le \mathbb{E}_A\left[\ \mathbb{I}\left(im \ge im_{\alpha,o}\right)\right] \\
&\quad \left(\text{if Eq. (14) holds, then equality holds}\right) \\
&= \mathbb{E}_A\left[\ \Pr\left(im \ge IM \mid EDP = edp_o, A = \alpha\right)\right] \\
&= \mathrm{P}\left(IM \le im \mid EDP = edp_o\right)
\end{aligned} \tag{15}$$

where A denotes a seismic event. Let us consider, for example, two IDA curves shown in Figure 1. The fragility value at im^* is of interest given the threshold level of edp_o. If one follows the exact fragility definition in Eq. (2), the fragility value is 0 (by counting the points satisfying $\{EDP \ge edp_o\}$ on the horizontal line of $IM=im^*$). However, if one follows the so-called IM-based definition in Eq. (12) (by counting the points satisfying $\{IM \le im^*\}$ on the vertical line of $EDP=edp_o$), the fragility value is estimated as 1/2, which is a clear overestimation. Therefore, for precise computation of fragility values, one may want to avoid the IM-based definition. It is also noted that the resulting error is reduced as a larger number of ground motion cases is considered or when a parametric (typically the lognormal CDF) function is fitted to the relevant data.

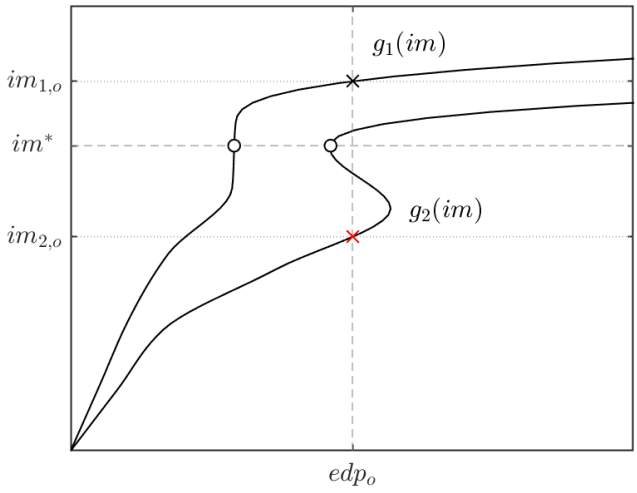


Figure 1. Example of IDA curves leading to fragility overestimation

2.6 *Extended fragility analysis*

Extended fragility analysis (Andriotis & Papakonstantinou, 2018) adopts and extends the theory of Generalized Linear Models (GML), often presented in the form of:

$$\mathbb{E}[Y] = g^{-1}(\mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}) \tag{16}$$

where $\mathbf{X}$ is the vector of input variables, $\boldsymbol{\beta}$ is the coefficient vector, and $g(\bullet)$ is some link function that relates the linear predictor and the output of interest. Particularly for a classification problem, where Y is a binary variable and its mean is bounded to the range of $E[Y] \in [0,1]$, the probit and logistic functions are common choices related to link functions (Calvin and Long, 1998).

Note that the formulation by Shinozuka *et al.* described in Section 2.2 is practically equivalent to a probit regression model. Similarly, extended fragility analysis introduces the logit link function, which, in the single IM case, is:

$$\ln\frac{p_f(im)}{1-p_f(im)} = \beta_0 + \beta_1 \ln im \tag{17}$$

where $\ln im$ is the predictor variable, and the left-hand side term is an inverse form of the logistic function or else a special case of the softmax function with a binary label. Rearranging, Eq. (17), the fragility form is derived:

$$p_f(im) = \frac{1}{1+e^{-(\beta_0+\beta_1 \ln im)}} \tag{18}$$

It has been shown by Andriotis and Papakonstantinou (2018) that the softmax function provides an accurate mathematical representation for fragility analysis under broad probabilistic assumptions. Further, it retains this property and smoothly extends to multivariate IM and multinomial DS fragility analysis, while bypassing important concerns discussed in the next section.

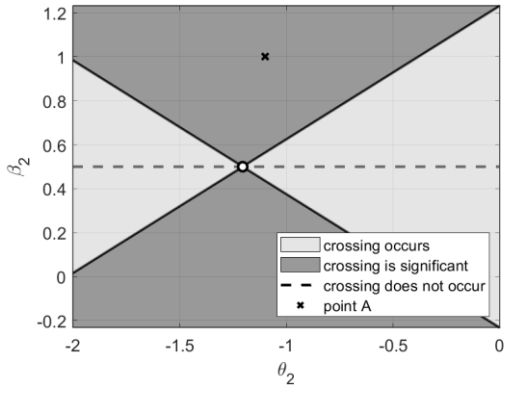
3 MULTIPLE DS AND MULTIVARIATE IM

3.1 *Multiple DS and crossing concerns*

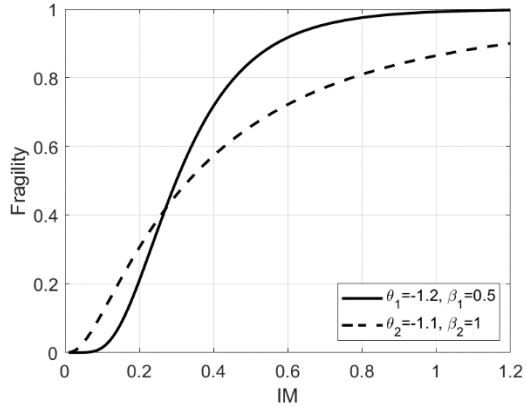
When multiple threshold levels are of interest, i.e., when DS is a nominal or ordinal variable with more than two states, the concern of crossings between fragility curves arises (Porter, 2007). As the intersection of two fragility curves necessarily implies that the probability of the system being in

a certain damage state is a negative value, it should be strictly prevented. However, when all the fragility curves maintain the same functional form, e.g., lognormal CDF, fragility curve crossings always occur unless all the curves are constrained to have the same dispersion. For example, consider two lognormal CDF functions. Provided that the first lognormal function has the parameters of θ_1=–1.2 and β_1=0.5, the domains, related to the parameters of the second function, where a crossing occurs and is significant are illustrated in Figure 2a. The fragility curves corresponding to point A in Figure 2a are illustrated in Figure 2b. Note that although this is a numerically possible outcome, it is physically unacceptable.

One way to avoid the crossings is to constrain the dispersion parameters to be common during the optimization (Shinozuka *et al*., 2003), which puts a strong limitation. On the other hand, the extended fragility method does not need to apply this restriction and successfully avoids crossings by allowing each fragility curve to have a different shape, not assigning a single CDF expression.



(a) Parameter space of the second function given the first fragility function parameters (θ_1, β_1)=(–1.2,0.5).



(b) The crossing of fragility curves at point A.

Figure 2. Crossing of fragility curves.

3.2 *Multiple DS with fixed dispersion*

As the *cloud analysis* already assumes the dispersion parameter α_{res} in its linear regression part to be the same over the range of thresholds, it circumvents crossings. Note that the logistic regression subpart of cloud analysis provides just a multiplication factor equally applied for different DS, and therefore it does not change the ordering between the fragility curves.

For the *Shinozuka et al.* method, in order to secure a common variance, the below form of fragility curves is introduced as an extension of Eq. (4):

$$p_{z\geq j}(im)=\Phi\left(\frac{\ln im-\theta_j}{\beta}\right) \tag{19}$$

where $p_{z\geq j}$ (•) indicates the fragility curve for the j-th damage state, while j=1 is an intact reference state. Note from Eq. (19), that the different fragility curves are merely horizontally shifted in the log-IM axis. The maximum likelihood estimator for the involved parameters is obtained from the following likelihood function:

$$L=\prod_{n=1}^{N}\prod_{j=1}^{J}p_{z_n=j}\left(im_n\right)^{II\{z_n=j\}} \tag{20}$$

where $II\{\bullet\}$ is again an indicator function, and $p_{zn=j}$ (•) denotes the probability that the DS of the n-th sample, z_n, belongs to the j-th damage state.

The variance constraint can also be *optionally* applied to *extended fragility analysis* of Eq. (18) as

$$p_{z\geq j}(im)=\frac{1}{1+e^{-(\beta_{0j}+\beta_1\ln im)}} \tag{21}$$

and substituted to the same likelihood as in Eq. (20). This method is referred to as the *ordinal* approach (Andriotis & Papakonstantinou, 2018).

3.3 *Multiple DS without crossings*

In its utmost generality, the *extended fragility* analysis, is not constrained to have the same variance among fragility curves nor a single CDF expression per fragility curve. This is defined by the *nominal* case, introduced by the following definition:

$$\ln\frac{p_{z=j}(im)}{p_{z=J}(im)}=\beta_{0j}+\beta_{1j}\ln im \tag{22}$$

where the *J*-th DS is referred to as a pivot DS. Eq. (22) can be equivalently written as:

$$p_{z=j}(im)=\begin{cases}\dfrac{e^{\beta_{0j}+\beta_{1j}\ln im}}{1+\sum_{i=1}^{J-1}e^{\beta_{0i}+\beta_{1i}\ln im}} & \text{for } j<J\\ \dfrac{1}{1+\sum_{i=1}^{J-1}e^{\beta_{0i}+\beta_{1i}\ln im}} & \text{for } j=J\end{cases} \tag{23}$$

Eq. (23) can be directly used for MLE optimization in Eq. (20). Once the parameters are identified, the fragility function is then obtained as the cumulative sum:

$$p_{z\geq j}(im)=\sum_{i=j}^{J}p_{z=i}(im) \tag{24}$$

Note that since $p_{z=j}(\bullet)$, $j=1,\ldots,J$, are always non-negative, Eq. (24) ensures that fragility curves of different DS levels do not intersect each other at any IM point.

3.4 *Multiple DS and multivariate IM*

Let us now consider multivariate IM cases where fragility curves become higher dimension fragility (*hyper-*)*surfaces* (Baker 2007, Ebrahimian *et al.* 2015). For the *cloud method*, the fragility function in Eq. (11) is generalized as:

$$\begin{aligned}p_{z\geq j}(\mathbf{im})=&\\ &\mathrm{P}(EDP\geq edp_j\,|\,\mathbf{im},NC)\mathrm{P}(NC\,|\,\mathbf{im})\\ &+\mathrm{P}(C\,|\,\mathbf{im})\end{aligned} \tag{25}$$

where **im** is now an M dimensional vector. Each logistic and linear regression part is parametrized as:

$$\mathrm{P}(C\,|\,IM=im)=\frac{1}{1+e^{-h_\beta(im)}} \tag{26}$$

and

$$\begin{aligned}P(EDP\geq edp_j\,|\,\mathrm{IM}=im,NC)&\\ =\Phi\left(\frac{edp_j-h_\alpha(im)}{\alpha_{res}}\right)&\end{aligned} \tag{27}$$

respectively, where:

$$h_\beta(\mathbf{im})=\beta_0+\beta_1\ln(im^1)+\cdots+\beta_M\ln(im^M) \tag{28a}$$

$$h_\alpha(\mathbf{im})=\alpha_0+\alpha_1\ln(im^1)+\cdots+\alpha_M\ln(im^M) \tag{28b}$$

Further, by introducing the methodologies of GML, *Shinozuka et al.* and *ordinal extended fragility* can be easily generalized to have multiple explanatory variables, as:

$$p_{z\geq j}(\mathbf{im})=\Phi\left(h_j(\mathbf{im})\right) \tag{29}$$

and

$$p_{z\geq j}(\mathbf{im})=\frac{1}{1+e^{-h_j(\mathbf{im})}} \tag{30}$$

respectively, where:

$$h_j(\mathbf{im})=\beta_{0j}+\beta_1\ln(im^1)+\cdots+\beta_M\ln(im^M) \tag{31}$$

Thus, by constraining the slope parameters to be shared among all DS, the two methods successfully avoid crossings between fragility surfaces. However, mere distinction with respect to intercept parameters reduces the flexibility of the models, as also described in Sections 3.1-3.2.

Finally, the *nominal* approach can similarly be generalized for multi-dimensional IM, as:

$$p_{z=j}(im)=\begin{cases}\dfrac{e^{h_j(\mathbf{im})}}{1+\sum_{i=1}^{J-1}e^{h_i(\mathbf{im})}} & \text{for } j<J\\ \dfrac{1}{1+\sum_{i=1}^{J-1}e^{h_i(\mathbf{im})}} & \text{for } j=J\end{cases} \tag{32}$$

where, in this case:

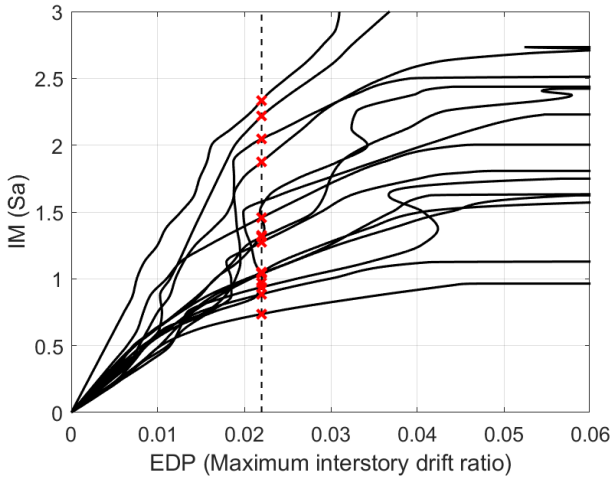
$$h_j(\mathbf{im})=\beta_{0j}+\beta_{1j}\ln(im^1)+\cdots+\beta_{Mj}\ln(im^M) \tag{33}$$

Note that the number of parameters in the nominal extension is $M(J-1)$, while in the ordinal case and probit regression $M+J-1$ parameters are required. The model flexibility can further be increased by introducing higher-order polynomial terms to Eq. (33) (Andriotis & Papakonstantinou, 2018). A final remark is that although this work describes the extended fragility analysis from a GLM perspective, the original work of Andriotis and Papakonstantinou (2018) used the more general form of softmax representation, that can additionally describe fragility functions when uncertain damage states are present.

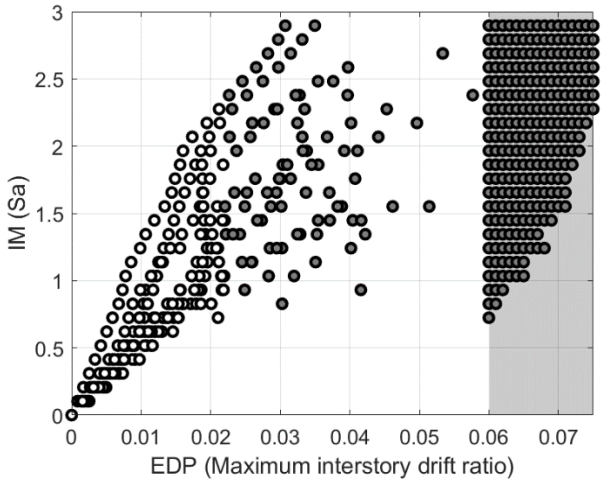
4 NUMERICAL EXAMPLES

The findings in this paper are further supported by numerical examples. A two-story moment resisting frame is considered, where beams and columns are respectively modeled by W24X131

and W27X102 sections. Analyses are conducted in OpenSees, version 3.2.1 (McKenna *et al*., 2010). The initial first and second period of the structure is 0.96 and 0.27 secs respectively, and the damping ratio is 0.05. The modified Ibarra-Krawinkler deterioration model is adopted to describe bilinear plastic hinges, and the details of the nonlinearity properties are retrieved from Eads (2010), while only the rotational capacity of the hinges is modified to θ_u=0.04 rad. Earthquake ground motions are generated from the site-specific ground motion generator introduced by Vlachos *et al*. (2016, 2018), with parameters in Table 1. Sixteen ground motions are selected from random realizations to demonstrate the findings. The spectral acceleration at 5% damping is used as the IM and the inter-story drift ratio as the EDP.



(a) IDA curves.



(b) Scattered data points.

Figure 3. Example of IDA curves.

Table 1. Parameters of site-specific earthquake model

Parameters	Value
fault distance	10-15 km
magnitude	6.0-8.0
ground velocity	212 m/s

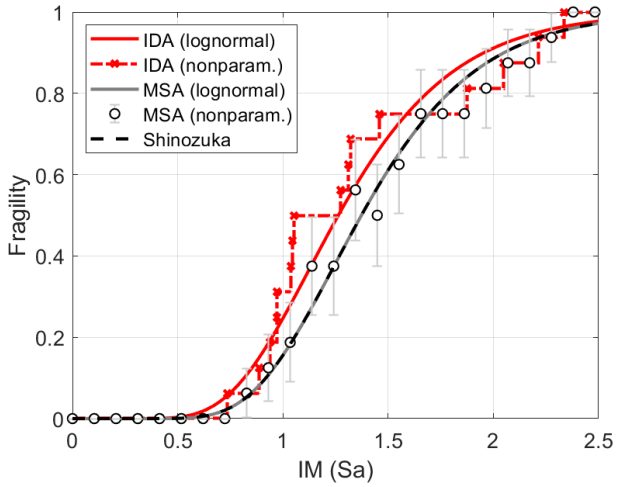
Figure 3a shows the obtained IDA curves, and Figure 3b the corresponding scattered data representation. To configure the same conditions for each method, the same dataset shown in Figure 3 is used for every fragility method. The fragility analysis results are plotted in Figure 4. It can be shown from Figure 4a that the parametrically fitted MSA and Shinozuka *et al*. are identical as they are mathematically equivalent. On the other hand, IDA overestimates the fragility curves due to the reversing behavior of IDA curves, as observed in Figure 3a. However, it should be noted again that although IDA is strictly an upper bound of the unbiased MSA estimator for the nonparametric results, this is not always true when it is fitted to a lognormal CDF. Instead, the curves from IDA and MSA can cross at some point. On the other hand, Figure 4b shows that Shinozuka *et al.* and extended fragility methods are in close agreement. It implies that the underlying functional form, i.e., lognormal and log-logistic CDFs form a close match with each other for the univariate IM and binomial DS case. However, it should be mentioned that as the problem dimensions increase, extended fragility allows a more flexible shape that does not rely on a single CDF expression over the IM-space. Finally, cloud analysis also shows some discrepancies arising from the violation of the linearity assumption in Eq. (9).

When a higher threshold level is of interest, i.e., when edp_o = 0.06, all five fragility curves give consistent results. The discrepancies in IDA and cloud method have been now removed, as (i) the obtained IDA curves were monotonic at the specific threshold level, and (ii) most of the failure cases are collapsed cases, enforcing the cloud analysis to be dominated by the logistic regression part.

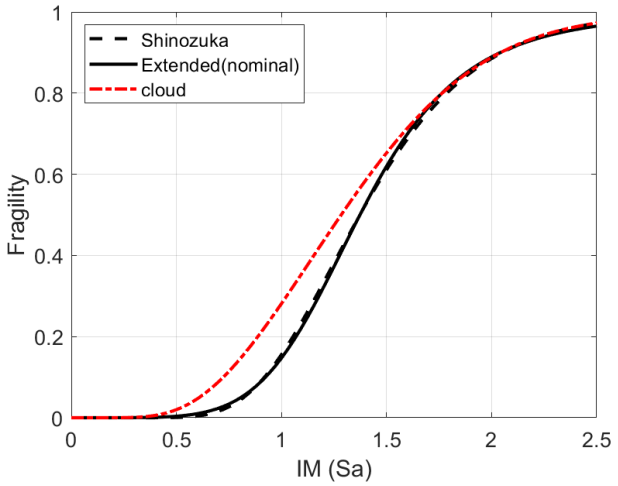
The problem is now further generalized to incorporate 2-dimensional IM. The significant earthquake duration, i.e., the time interval within an earthquake event of which Arias intensity lies between 5% to 95% of its final value, is added as an IM, and 50 ground motions are randomly

generated for this analysis. One can notice from Figures 5a and 5b that Shinozuka *et al*. and extended fragility (ordinal) methods are in good agreement as they are both constrained to have common variances for different threshold levels. On the other hand, cloud analysis results in Figure 5c show a large discrepancy due to the linear regression implications. Nominal extended fragility method results are illustrated in Figure 5d. It is shown that the surface for the most severe DS (the lowest surface) exhibits some differences with Figures 5a and 5b. By allowing a higher level of flexibility, the nominal extended fragility method is capable of describing a more precise shape of the fragility surface.

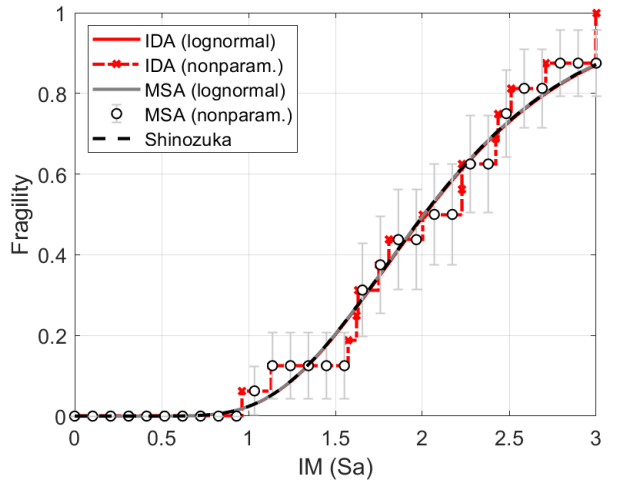
A third IM is further introduced, which is the ratio between the first and second mode spectral acceleration. Figure 6 presents the IM dataset used for fragility analysis. Instead of describing the fragility 3D-surface, the separation boundaries (Andriotis & Papakonstantinou, 2018) between neighboring DS are now illustrated along with the sample points in Figure 6. The described different attributes and constraints between ordinal and nominal fragility analyses are clearly showcased in the figures.



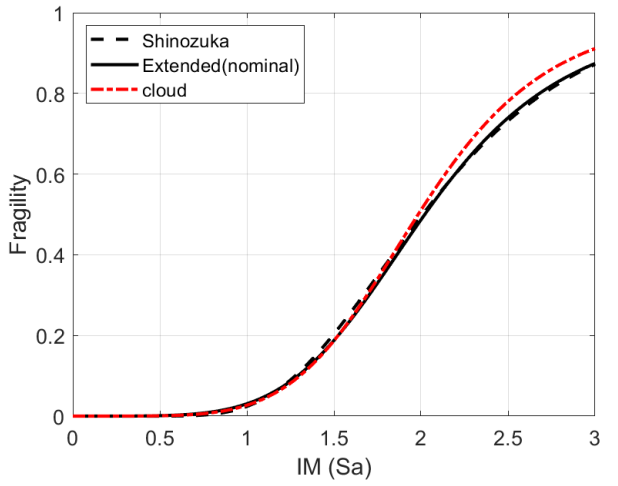
(a) IDA, MSA, Shinozuka *et al*. (edp_o=0.022).



(b) Shinozuka *et al*., extended, cloud (edp_o=0.022).

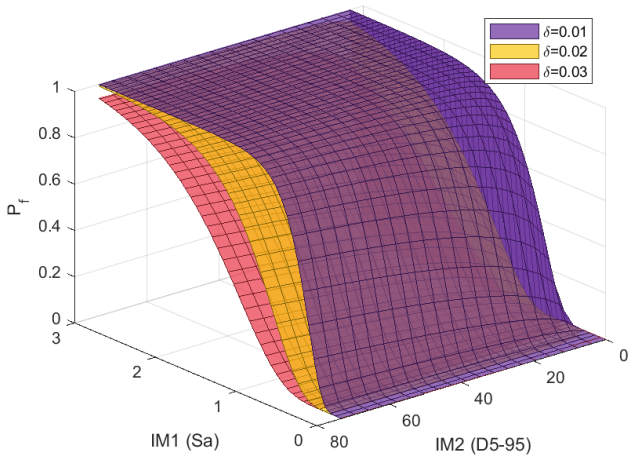


(c) IDA, MSA, Shinozuka *et al*. (edp_o=0.06).

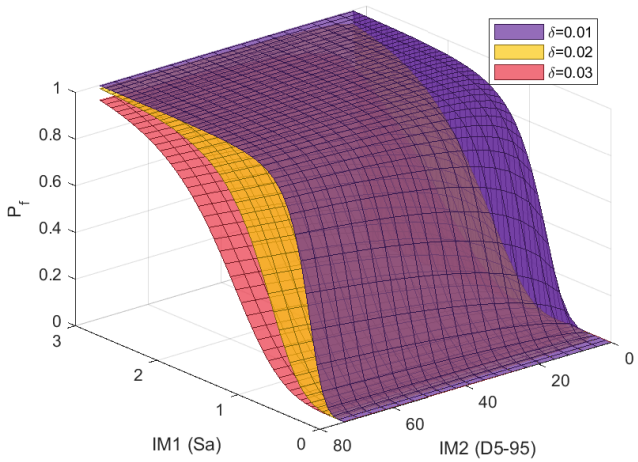


(d) Shinozuka *et al*., extended, cloud (edp_o=0.06).

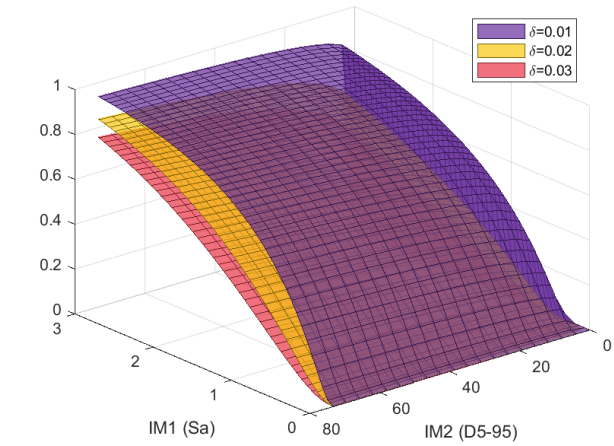
Figure 4. Fragility curves by different methods.



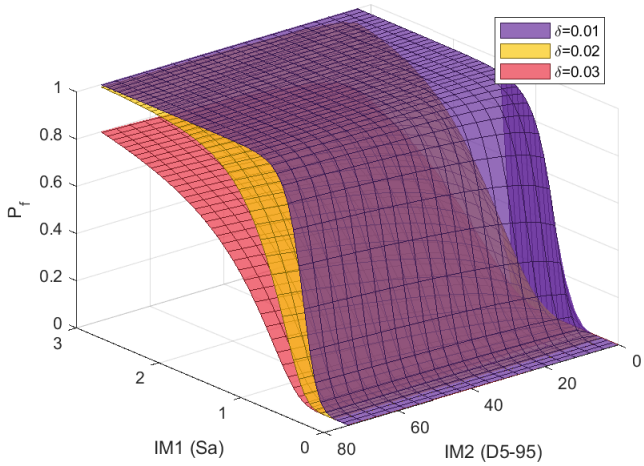
(a) Shinozuka et al.



(b) Extended fragility (Ordinal)

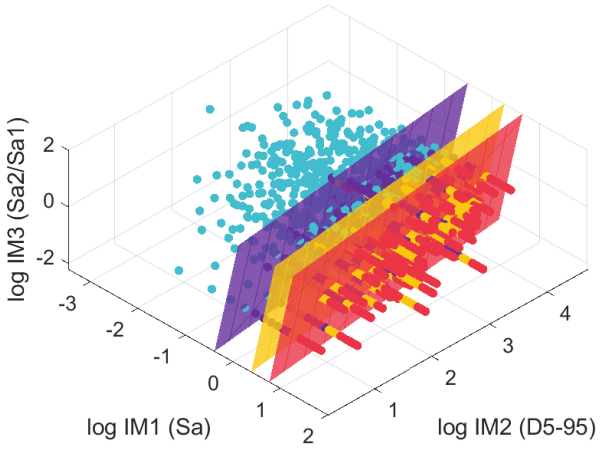


(c) Cloud method

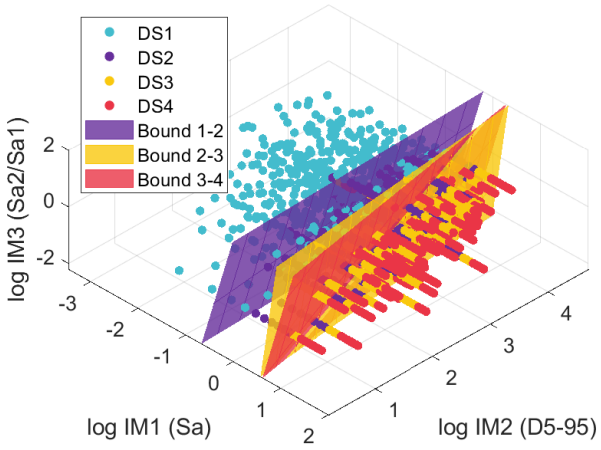


(d) Extended fragility (Nominal)

Figure 5. Fragility surfaces by different methods.



(a) Ordinal



(b) Nominal

Figure 6. Separation boundaries obtained by extended fragility methods.

5 CONCLUSIONS

This paper provides a mathematical appraisal and overview for different fragility analysis methods. The well-known approaches of multiple-stripe analysis (MSA), incremental dynamic analysis (IDA), and cloud analysis are examined, along with the classical Bernoulli-link model and the recently developed extended fragility framework. A unified viewpoint is presented, and mathematical proofs that discuss the eligibility of and interconnections between the methods are provided. It is shown that MSA is a subclass of Shinozuka *et al*. and IDA is an upper bound of the unbiased fragility estimator. Shinozuka *et al*. and extended fragility analysis can be viewed through the framework of generalized linear models, but the extended fragility has also capabilities beyond this framework as far as function choices, versatility, and extensibility are concerned.

Extended fragility analysis, cloud analysis, and Shinozuka *et al*. are then studied for multivariate IM and multiple DS cases, where only the extended fragility approach provides the most flexible descriptions without allowing crossings between the fragility curves.

The unique integrative perspective in this study further supports several new research and practical directions, such as state-dependent transitions in time and hidden condition states. Overall, this work aims at providing several insights, enabling efficient, accurate, and consistent probabilistic estimations of structural conditions, in earthquake engineering and beyond. Finally, it is remarked that a more complete study of this topic is in preparation for publication (Papakonstantinou *et al.*, in preparation).

ACKNOWLEDGEMENTS

Dr. Papakonstantinou acknowledges the support of the U.S. National Science Foundation under CAREER Grant No. 1751941. Dr. Andriotis also acknowledges the support of the TU Delft AI Labs program.

REFERENCES

Andriotis C.P., Papakonstantinou K.G. 2018. Extended and generalized fragility functions. *Journal of Engineering Mechanics*, 144(9), 04018087.

Bakalis, K., & Vamvatsikos, D. 2018. Seismic fragility functions via nonlinear response history analysis. *Journal of Structural Engineering*, 144(10), 04018181.

Baker, J.W. 2015. Efficient analytical fragility function fitting using dynamic structural analysis. *Earthquake Spectra*, 31(1), 579-599.

Baker, J.W. 2007. Probabilistic structural response assessment using vector-valued intensity measures. *Earthquake Engineering & Structural Dynamics*, 36(13), 1861-1883.

Calvin J.A., Long J.S. 1998. Regression models for categorical and limited dependent variables. *Technometrics*. 40:80–1.

Eads. L. 2010. *Dynamic analysis of 2-story moment frame.* English, Stanford University, Available: http://opensees.berkeley.edu/wiki/index.php/Dynamic_Analysis_of_2-Story_Moment_Frame

Ebrahimian, H., Jalayer, F., Lucchini, A., Mollaioli, F., & Manfredi, G. 2015. Preliminary ranking of alternative scalar and vector intensity measures of ground shaking. *Bulletin of Earthquake Engineering*, 13(10), 2805-2840.

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

Jalayer F, Ebrahimian H, Miano A, Manfredi G & Sezen H. 2017. Analytical fragility assessment using un-scaled ground motion records. *Earthquake Engineering and Structural Dynamics,* 46(15):2639-2663.

McKenna, F., Scott, M.H., & Fenves, G.L. 2010. Nonlinear finite-element analysis software architecture using object composition. *Journal of Computing in Civil Engineering*, 24(1), 95-107.

Miano A, Jalayer F, Ebrahimian H & Prota A. 2018. Cloud to IDA: Efficient fragility assessment with limited scaling. *Earthquake Engineering & Structural Dynamics*, 47(5):1124-1147.

Papakonstantinou, K.G., Yi, S., Andriotis, C.P., & Song, J. Mathematical foundations of fragility functions. *Manuscript in preparation.*

Porter, K., Kennedy, R., & Bachman, R. 2007. Creating fragility functions for performance-based earthquake engineering. *Earthquake Spectra*, 23(2), 471-489.

Shinozuka M, Feng MQ, Kim H, Uzawa T, Ueda T 2003. *Statistical analysis of fragility curves.* MCEER Technic al Report.

Shinozuka, M., Feng, M. Q., Lee, J., & Naganuma, T. 2000. Statistical analysis of fragility curves. *Journal of Engineering Mechanics*, 126(12), 1224-1231.

Vamvatsikos, D., & Cornell, C.A. 2002. Incremental dynamic analysis. *Earthquake Engineering & Structural Dynamics,* 31(3), 491-514.

Vamvatsikos, D., & Cornell, C.A. 2004. Applied incremental dynamic analysis. *Earthquake Spectra*, 20(2), 523-553.

Vlachos, C., Papakonstantinou, K.G., & Deodatis, G. 2016. A multi-modal analytical non-stationary spectral model for characterization and stochastic simulation of earthquake ground motions. *Soil Dynamics and Earthquake Engineering*, 80, 177-191.

Vlachos, C., Papakonstantinou, K.G., & Deodatis, G. 2018. Predictive model for site specific simulation of ground motions based on earthquake scenarios. *Earthquake Engineering & Structural Dynamics*, 47(1), 195-218.