ERROR ESTIMATION OF CLOSED-FORM SOLUTION FOR ANNUAL RATE OF STRUCTURAL COLLAPSE

Brendon A Bradley*, Rajesh P Dhakal

Department of Civil Engineering, University of Canterbury, Private Bag 4800, Christchurch 8020, New Zealand

*Corresponding author: Ph +64-3-366 7001 ext 7333; Fax: +64-3-364 2758;

Email: bab54@student.canterbury.ac.nz

ABSTRACT

With the increasing emphasis of performance-based earthquake engineering (PBEE) in the engineering community, several investigations have been presented outlining simplified approaches suitable for performance-based seismic design (PBSD). Central to most of these PBSD approaches is the use of closed-form analytical solutions to the probabilistic integral equations representing the rate of exceedance of key performance measures. Situations where such closed-form solutions are not appropriate primarily relate to the problem of extrapolation outside of the region in which parameters of the closed-form solution are fit. This study presents a critical review of the closed form solution for the annual rate of structural collapse. The closed form solution requires the assumptions of lognormality of the collapse fragility and power model form of the ground motion hazard, of which the latter is more significant regarding the error of the closed-form solution. Via a parametric study, the key variables contributing to the error between the closed-form solution and solution via numerical integration are illustrated. As these key variables can not be easily measured it casts doubt on the use of such closed-form solutions in future PBSD, especially considering the simple and efficient nature of using direct numerical integration to obtain the solution.

KEYWORDS

Performance-based seismic design (PBSD), performance-based earthquake engineering (PBEE), ground motion hazard, annual rate of collapse, deaggregation.

INTRODUCTION

Performance-based earthquake engineering (PBEE) and performance-based seismic design (PBSD) concepts are growing in popularity amongst the engineering community because of their consistent nature with respect to the socio-economic aims of seismic design. PBEE and PBSD also allow for incorporation of the uncertainties in all aspects of seismic design and assessment. The growing importance of PBEE and PBSD is illustrated by its inclusion in recent significant documents [1-5] .

Typical key performance measures in PBEE include the annual rate of exceedance of a given level of demand or financial loss, and the annual rate of structural collapse. The direct incorporation of uncertainties in the aforementioned performance measures results in an integral equation, which is an application of the total probability theorem. In such equations, a cumulative density function (CDF) is integrated over all intensities with the ground motion hazard curve for a specific site.

A key concept advocated by researchers in this area is that for PBSD to be accepted in design, simplified methods must be available which allow reasonably accurate evaluations to be made based on sound underlying assumptions. For the aforementioned key performance measures, numerous references are available for 'closed-form' analytical solutions. The first closed-form solutions were published for the demand hazard in References [6, 7], and using similar assumptions, annual frequencies of limit state exceedance and structural collapse can also be computed [8-10].

Such closed-form solutions have been used extensively since their development. Cornell *et al*. [9] used the closed-form drift hazard solution in a load and resistance factor design (LRFD) approach, which is implemented in FEMA-350 [3]. Mackie and Stojadinovic [11] used closed-form solutions for damage and loss limit states to propose a PBSD approach for bridges. Zareian and Krawinkler [10] used the closed form solution for the annual rate of

collapse, to propose a PBSD methodology considering structural collapse. The above three implementations also separate epistemic and aleatory uncertainties in the structural response and use the mean ground motion hazard curve. These two treatments allow the determination of the mean annual rate of exceedance of a particular performance measure with a specified level of confidence.

The closed-form solution for the annual exceedance rate of demand (i.e. demand hazard) is based on the following three assumptions: (i) the ground motion hazard curve is approximated by a linear line in log-log space; (ii) the median demand given intensity is a linear function in log-log space; and (iii) the demand given intensity distribution is assumed lognormal with constant logarithmic standard deviation (herein referred to as the 'dispersion') over the range of intensity that is of interest. Because these assumptions are made only in the regions of interest of the relationships, then the resulting closed-form solution may be considered as a 'local approximation' of the key performance measure around the region of interest. For example, it is stated in Kennedy and Short [6] that "over any ten-fold difference in exceedance probabilities, such hazard curves may be approximated by the PSDA analytical equation".

Aslani and Miranda [12] compared the closed-form solution for the demand hazard with that obtained by direct numerical integration using parametric relationships for the mean and dispersion of the demand given intensity relationships. They illustrated the resulting error in the demand hazard curves due to each of the three aforementioned assumptions required in the closed-form solution becomes significant as the demand levels become significantly different from those which the parameters were fit too.

Recently, Bradley *et al*. [13] proposed a 'hyperbolic' parametric equation to represent the ground motion hazard which is significantly more accurate over a larger range of exceedance frequencies than the power-model equation used to obtain the closed form

solution for the demand hazard (i.e. [6, 7]). It was then illustrated how a semi-analytical solution for the demand hazard could be obtained using the 'hyperbolic' hazard model. This work offered a potential solution to the problem of 'extrapolation' of the local approximation of the closed-form demand hazard solution to a larger range of exceedance frequencies. However, in computing the exceedance rate of a single value of demand, the semi-analytical solution of Bradley *et al*. [13] and the closed-form solution using the power-model equation are identical. Also, the semi-analytical solution given by Bradley *et al*. [13] still requires the two assumptions for the demand given intensity relationship which also introduce some extrapolation error [12].

From the above discussions, it is clear that the criticism of the closed-form solutions is primarily due to their inability to accurately extrapolate outside the immediate range over which the parametric relationships are fit. This paper investigates the error in the closed-form solution for the annual rate of structural collapse (collapse hazard), which does not suffer from the problems of extrapolation as the demand hazard mentioned above; implications related to the demand hazard are also briefly addressed. Deaggregation [14-16] of the integral equation is used to determine the regions of ground motion intensity which significantly contribute to the numerical value of the collapse hazard. Via a parametric study, key features of the integral equations that contribute to the error between the closed-form and exact numerical solutions are identified. Various means of fitting the power-model equation to the ground motion hazard data are discussed in light of the resulting errors in the parametric study.

CLOSED-FORM SOLUTION FOR THE ANNUAL RATE OF STRUCTURAL COLLAPSE

Firstly, the Pacific Earthquake Engineering Research (PEER) Centre PBEE framework terminology is adopted herein. Therefore, seismic demand is referred to as an engineering

demand parameter (*EDP*), and ground motion intensity as an intensity measure (*IM*). The 5% damped elastic spectral acceleration at the fundamental period of the structure ($Sa(T_1,5\%)$, herein *Sa* for brevity) is used as the *IM*.

The annual rate of structural collapse (collapse hazard) can be obtained by integrating (over the entire range of ground motion intensity) the conditional probability of collapse for a given level of intensity with the incremental probability of occurrence of that ground motion intensity. The mathematical formulation of the collapse hazard is given in Equation (1), which is an application of the Total Probability Theorem [17]:

$$\lambda_C = \int_0^\infty P\left(C \middle| IM = im\right) \left| \frac{d\lambda(IM > im)}{dIM} \right| dIM \tag{1}$$

where λ_C = the annual rate of collapse; $P(C \mid IM = im)$ = the conditional probability of collapse given $IM = im$ (collapse fragility curve); and $\lambda(IM > im)$ = the annual rate of exceedance of $IM = im$ (ground motion hazard) at the site. The absolute value signs around the derivative of the ground motion hazard are used as its value is negative.

In order to obtain a closed-form solution of Equation (1), several simplifying assumptions are required. Firstly, the intensity at which collapse is observed to occur is assumed to be of the form given in Equation (2):

$$\ln\left(IM \mid C\right) = \ln(\eta_Z) + \ln(\varepsilon_{UZ}) + \ln(\varepsilon_{RZ}) \tag{2}$$

where η_Z = the median *IM* causing collapse; and $\ln\left(\varepsilon_{RZ}\right)$ and $\ln\left(\varepsilon_{UZ}\right)$ are aleatory and epistemic uncertainties having a normal distribution with zero mean and standard deviations of β_{RZ} and β_{UZ}, respectively. Equation (2) results in a collapse fragility curve (due to aleatory randomness) which has a cumulative lognormal distribution, and η_Z also having a lognormal distribution.

The ground motion hazard is also assumed to have a linear form in log-log space given by Equation (3):

$$\ln\left(\lambda(IM)\right) = \ln(k_0) - k.\ln(IM) + \ln(\varepsilon_{UIM}) \tag{3}$$

where k_0 and k are constants fitted to the ground motion hazard in the region of interest [13], and $\ln(\varepsilon_{UIM})$ is a normal random variable with zero mean and standard deviation β_{UIM}, representing epistemic uncertainty in the ground motion hazard. Hence, the mean of Equation (3) is $\overline{\lambda(IM)} = k_0 IM^{-k}$. One further assumption is that $\ln(\varepsilon_{UZ})$ and $\ln(\varepsilon_{UIM})$ are independent of each other, but within each random variable there is a perfect correlation at various levels of intensity (e.g. $\ln(\varepsilon_{UZ})$ is perfectly correlated to itself at various levels of intensity).

Based on the aforementioned assumptions, the evaluation of Equation (1) using integration by parts leads to the following closed-form solution for the mean collapse hazard (See Jalayer [18] for details on a similar process to obtain the demand hazard):

$$E[\lambda_C] = k_0 \eta_Z^{-k} \exp\left[\frac{1}{2} k^2 \left(\beta_{UZ}^2 + \beta_{RZ}^2\right)\right] \tag{4}$$

Furthermore, λ_C is a lognormal random variable with dispersion:

$$\sigma_{\ln \lambda_C} = \sqrt{\beta_{UIM}^2 + k^2 \beta_{UZ}^2} \tag{5}$$

Equation 4 indicates that the expected value of the annual rate of collapse can be obtained from the annual frequency of exceedance of the median *IM* value causing collapse, η_Z, and then a multiplying factor (the exponential term) which represents the effect of uncertainty on the annual frequency of structural collapse. This factor indicates that increasing the uncertainty in the collapse fragility curve and the log-log slope of the ground motion hazard curve, increases the expected frequency of collapse. In particular, it is noted that while increasing the dispersion of the collapse fragility curves increases the probability of collapse at *IM* values lower than the median *IM* but reduces the probability of collapse at *IM* values larger than the median *IM*, it is the small IM values which occur significantly more frequently.

Figure 1a gives a typical probabilistic seismic demand analysis (PSDA) plot which has

been derived via Incremental Dynamic Analysis (IDA) [19] of a single-degree-of-freedom (SDOF) model of a New Zealand bridge pier. The SDOF model uses a lumped plasticity (frame) element with the modified Takeda hysteresis having both strength and stiffness degradation. Further details on the bridge structure and its modelling can be found in Reference [20]. Each of the lines in Figure 1a represent the result of an individual record scaled over a range of *IM*, and the dots at the end of the lines represent the projection (to the right boundary of the figure) of the last *IM* value before structural collapse was observed. Structural collapse is considered as the limit state of global sidesway instability (indicated numerically by non-convergence of the analysis). Global collapse associated with loss of vertical carrying capacity (due to axial and/or shear failures) is not considered here due to the lack of reliable analysis tools for capturing such phenomena [10]. Others have also defined global collapse when the slope of the tangent of the IDA curve drops below 20% of the initial tangent [3, 21], but this was not done here. Based on the sample mean and standard deviation of the *IM*'s causing collapse, a lognormal distribution of collapse given *IM*, can be defined, which is also shown in Figure 1a. Figure 1b gives a typical comparison between the seismic hazard curve for Wellington, New Zealand, and the approximation of the power-model (Equation (3)), fitted tangentially to the median *IM* causing collapse of the bridge structure considered.

SOURCES OF ERROR IN COLLAPSE HAZARD CLOSED-FORM SOLUTION

Firstly, discussions are restricted to the error associated with the expected value of the collapse hazard (i.e. Equation (4)), and consider only one source of uncertainty in the collapse fragility curve. This uncertainty may be solely aleatory, or a square-root sum squares (SRSS) [10] combination of both aleatory and epistemic uncertainties.

Equation (1) illustrates that the collapse hazard is a function of both the collapse fragility curve and the derivative of the ground motion hazard curve. Figure 2 gives a comparison of the lognormal collapse fragility curve and the empirical CDF based on the IDA data in Figure 1a. It can be seen that the typical [10] lognormal approximation is acceptable for this data, based on Kolmogorov-Smirnov (K-S) goodness of fit test [17]. Various other studies have illustrated that this assumption is adequate and it has been used via direct numerical integration with the full representation of the seismic hazard of the site [10, 12]. A non-parametric form of the collapse fragility can be used, however care should be taken to ensure that enough ground motions are used such that the annual frequency of structural collapse is not sensitive to the 'steps' in the empirical CDF. As an alternative to developing collapse fragility curves via IDA data, various data is available for collapse capacities for generic moment resisting frames and shear walls, which are useful for preliminary design assessments [22].

Based on the above discussion as the lognormal assumption for the collapse fragility curve is adequate, it will be shown that the most restrictive assumption in order to derive Equation (4) is the power-model approximation of the ground motion hazard curve. The power-model therefore assumes that the ground motion hazard is linear in log-log space which is considered as a 'local approximation'. The potential error comes from the fact that as Equation (1) involves integration over the entire range of *IM*, the power-model solution will potentially inaccurately approximate the likelihood of ground motions of $IM = im$ occurring over a large range of *IM*. This potential inaccurate approximation is due to the typical 'concave from below' shape of ground motion hazard curves in log-log space [13], compared with the linear (in log-log space) curve of the power-model.

As the power-model assumes that the ground motion hazard is linear in log-log space, the error will likely be a function of the 'curvature' of the hazard curve. Here, 'curvature' (ϕ)

is defined as the second derivative of the ground motion hazard curve in log-log space (i.e. the rate of change of the tangential slope, k). Because Equation (1) combines the ground motion hazard curve with the cumulative probability of collapse, the major contribution to the integral will occur from ground motion intensities around the central *IM* value causing structural collapse, η_Z. For example, in the limiting deterministic case (when there is no uncertainty), only η_Z is used to evaluate Equation (1). The range of *IM* values that significantly contribute to the integral (and hence the error in the closed-form solution) will therefore be a function of the likelihood of these *IM* values causing collapse to occur, $P(C \mid IM = im)$. Hence, any error in the closed-form solution (Equation (4)) will also be a function of the dispersion in the collapse fragility curve (herein denoted simply as β).

PARAMETRIC STUDY ON ERROR IN CLOSED-FORM SOLUTION USING A TANGENT-FIT TO HAZARD DATA

To investigate the effects of curvature, ϕ, and dispersion, β on the error in the closed-form solution, a parametric study was carried out which is described in the following paragraphs. For brevity, the term 'hazard' will be used in reference to 'ground motion hazard'. Note that both the closed-form solution and 'exact' numerical integration solution compared here use the lognormal assumption for the collapse fragility curve (i.e. not the raw data depicted in Figure 2). Therefore, differences between the outcomes of these two approaches are solely due to the representation of the ground motion hazard curve.

To obtain an estimate of the curvature of the hazard curve around the region of interest, the parametric form for the ground motion hazard model proposed by Bradley *et al*. [13] is used, which is given by:

$$E[\ln(\lambda(IM))] = \ln(\lambda_{asy}) + \alpha\left[\ln\left(IM/IM_{asy}\right)\right]^{-1} \tag{6}$$

where λ_{asy}, IM_{asy}, and α are constants to be fit by nonlinear regression. For the above parametric form the curvature at a given point can be found from:

$$\phi = \frac{\partial^2[\ln(\lambda(IM))]}{\partial[\ln(IM)]^2} = 2\alpha\left[\ln\left(\frac{IM}{IM_{asy}}\right)\right]^{-3} = \frac{2}{\alpha^2}\left[\ln\left(\frac{\lambda(IM)}{\lambda_{asy}}\right)\right]^3 \quad (7)$$

where the central and right-hand side algebraic expressions are the curvature as a function of IM and λ, respectively. Herein, unless otherwise stated, ϕ is calculated at the median IM corresponding to the collapse, η_Z.

To account for the fact that this simple definition of curvature will not be an exact measure of the error, five hazard curves for the major centres in New Zealand [23] were used. The hazard curves for these five regions, along with their curvatures as a function of rate of exceedance are presented in Figure 3a and Figure 3b, respectively. It can be seen that these hazard curves represent a wide range of site seismicity, from low in Auckland, to high in Otira. Figure 3b shows that the curvature of the hazard curves increases as the rate of exceedance reduces. It is also interesting to note that the curvature of the hazard curves is not directly related to the seismicity of the site. For example, the Christchurch hazard has a far larger curvature than the Wellington hazard, despite the Wellington hazard having a larger seismicity. A similar comparison between the Auckland and Dunedin hazards can also be made. As it will be shown later, the error in the closed-form solution increases as the curvature of the ground motion hazard increases. This indicates that the error is not directly related to the seismicity of the site.

In order to illustrate that the error in the closed-form solution (Equation (4)) is a function of both ϕ and β the concept of deaggregation [14-16] is used. Deaggregation allows the contribution of different values of the integrand to the integral to be graphically illustrated. Figure 4 shows four deaggregation plots of Equation (1) using both the 'exact' numerical solution and the closed-form solution, where the parameters of the power-model of the ground

motion hazard (Equation (3)) have been obtained by fitting the model tangentially to the raw hazard data at $IM = \eta_Z$. In these figures, the Christchurch ground motion hazard curve (which is of moderate seismicity) was used. Two frequencies for the median *IM* causing collapse and two values of the dispersion were considered. The frequencies for the median *IM* causing collapse considered were $\lambda = 2.1\text{x}10^{-3}$ and $\lambda = 1\text{x}10^{-4}$ (i.e. from the ground motion hazard curve, the median *IM* causing collapse, η_Z, has these exceedance frequencies). These two frequencies represent the upper and lower ranges of likely collapse frequencies. For example, non-ductile flexure-shear critical structures typically have an annual rate of collapse which can be greater than $2.1\text{x}10^{-3}$ (e.g. [24]), while for current code-conforming structures the collapse hazard is typically lower than $1\text{x}10^{-4}$ (e.g. [25]).

The first dispersion value used was $\beta = 0.3$. This dispersion value would typically occur for 'efficient' [8, 26] *IM* such as the inelastic spectral displacement, S_{di}, proposed by Tothong and Luco [27]. The second value of $\beta = 0.5$ was used as a value representative of dispersions due to a relatively inefficient *IM* (such as elastic spectral acceleration, S_a, which is the most commonly used *IM*). For example, although not explicitly mentioned, the dispersion (due to aleatory uncertainty) in the collapse fragility (using the first mode spectral acceleration as the *IM*) given in Reference [10] is approximately 0.42. Other cases where a large dispersion may be measured could be where: (i) several designs are to be compared, which do not have the same characteristics (e.g. fundamental period), in which case the use of a structure-dependent *IM*'s (such as S_a) may not be appropriate (e.g. Reference [11] gives 13 dispersion values ranging from 0.33 to 0.56 for simple bridge structures using *IM* = *PGV*); (ii) higher-mode effects are important (e.g. in flexible structures an *IM* such as S_a may not accurately predict a multi-mode dominated response [28]); and (iii) near-fault velocity-pulse effects [28].

It can be seen in Figure 4 that as the curvature and dispersion increase so does the error between the closed-form solution and the 'exact' solution using numerical integration. Here,

the error has been represented in the form of an error ratio, defined as:

$$E_{ratio} = \frac{\lambda_{C,approx}}{\lambda_{C,exact}} \tag{8}$$

where $\lambda_{C,approx}$ = the closed-form solution (Equation (4)); and $\lambda_{C,exact}$ = the 'exact' numerical solution of Equation (1). Figure 4a-Figure 4c therefore have errors of 16%, 60%, and 77%, respectively, while Figure 4d has a 7-fold (700%) error. It is also observed that the integration error contributed by *IM* values larger than the median *IM* causing collapse, η_Z, is negligible compared to the error contributed by *IM* values below η_Z. This consistent nature of the error in the closed-form solution potentially allows other means of fitting the ground motion power-model which is discussed in the following section.

Based on typical values for the dispersion observed in the literature [7-13, 18-20, 26-30] and exceedance rates of collapse that could occur for a wide range of structures [24, 25], a parametric study was performed using $\beta = 0.2\text{-}0.6$ and $\lambda = 10^{-2}\text{-}10^{-5}$. The results of the parametric study are presented graphically in Figure 5. Figure 5a shows the error ratios (as defined in Equation 8) for $\beta = 0.2$ and 0.3. The dashed lines surrounding the data points are used to clearly define the data points for each β value. The relatively small scatter between the data points for the five different hazard curves indicates that β and ϕ capture the salient features of the error between the closed-form solution and the 'exact' numerical solution. Figure 5b shows the results for $\beta = 0.4 - 0.6$. Again, the dashed lines are used to distinguish between different β values. It is obvious from both figures that the variation in error between the results for different hazard curves increases as β increases.

To give a practical viewpoint of Figure 5, consider the use of the closed-form solution with $\beta = 0.42$. This value of β is that (approximately) obtained in Reference [10], and is below the median of the β values used in Reference [11]. Assume that the structure is designed to current ductile design philosophy and has a fundamental period of T=1.5s (i.e. so

that the hazard curves of Figure 3a are used), and a median collapse intensity, η_Z = 1.4g (this is slightly less than $\eta_Z \sim$ 1.75g used in Reference [10], in which the structure had a period of T = 1.2s). Based on the results of Figure 5b the error ratios for the Wellington and Otira sites would be approximately 3.1 and 10.1, respectively. This means that if the collapse rate for Otira was found (using the closed-form solution) to be on the order of $\lambda = 10^{-4}$, then its actual value is likely to be in the region of $\lambda = 10^{-5}$. Note also, that the value of $\beta = 0.42$ represents aleatory uncertainty only. If epistemic uncertainty (which is typically in the region of 0.4-0.45 [10, 30] is also included in an SRSS form, then $\beta \sim 0.6$ and the error ratio will be in excess of 20. Such large errors defeat the purpose of using a probabilistic-based measure of performance.

ALTERNATIVE NON-TANGENT POWER-MODEL FITS TO GROUND MOTION HAZARD

The deaggregation results of Figure 5 illustrated that using a tangent based fit of the hazard curve to determine k results in significant over-approximation of the contribution of ground motions with $IM < \eta_Z$. This occurs because the log-log slope of the hazard curve, k, is too large over the region $IM < \eta_Z$. Therefore, a reduction in the value of k will likely reduce such an over-approximation. Such non-tangent methods have been suggested previously by others. For example, when computing the demand hazard around the design basis earthquake (DBE) and maximum considered earthquake (MCE) frequency region, Jalayer [18] suggested using fitting the power-model hazard as a secant through the DBE and MCE points of the ground motion hazard.

In this work, several alternative methods of fitting k were investigated, which include some of the following: (i) multiply the tangent-based fit of k by some constant; (ii) fit k tangential to hazard curve at some rate less than η_Z; (iii) use a secant-based fit of k between

two points either side of η_Z; and (iv) use regression over some region of the ground motion hazard to determine the power-model parameters. Table 1 gives a summary of the resulting error ratios for a selection of the different fitting methods used, for the Christchurch hazard. For example, using the secant-based fit with one point at $IM = \eta_Z$, and the other at a value of IM which has rate of exceedance equal to ten times that of η_Z (first row for base case (iii) in Table 1) resulted in relatively accurate (compared to the tangent-fit) results over the wide range of values (and different hazard curves) used in the parametric study. Figure 6a and Figure 6b give the deaggregation plots obtained using the secant-based fitting of k at $IM = \eta_Z$ and $IM_{10\lambda}$, which are for the same (φ, β) scenarios as Figure 4b and Figure 4d which used the tangential fit of k. The two vertical dashed lines in Figure 6a and Figure 6b show the IM values through which the secant-fit was performed. In particular, for β = 0.3 and $\lambda = 10^{-4}$ (Figure 6a), the error ratio for the secant-based fit is 2% (Figure 6a) compared to the 77% error using the tangent-based fit (Figure 4b). From the discrepancies between the numerical and closed-form solutions relative to the points where the secant-fit was performed, it becomes obvious that for this type of fitting, the closed-form solution under-predicts the contribution from ground motion intensities with $IM > IM_{10\lambda}$ and over-predicts the contribution of ground motion intensities with (approximately) $IM < IM_{10\lambda}$. Hence, the accuracy reflected in the error ratio of 1.02 is the result of 'subtractive cancellation', that is errors in one region are negated by errors (of opposite nature) in another region. Obviously, over a large range of β and φ values it is unlikely that such 'subtractive cancellation' will consistently occur. This is illustrated in Figure 6b, where for β = 0.5, $\lambda = 10^{-4}$, the error ratio is 1.86; still a significant reduction however compared to the 7-fold error using the tangent-based solution (Figure 4d).

Several other fitting methods such as 'k=0.75k_t' and '$IM_{0.5\lambda}$' from Table 1 appear to be more accurate, particularly at large values of β and φ. This however results from the

aforementioned 'subtractive cancellation', and these results significantly under-predict the exact value for small β and φ, yielding error ratios of 0.93 and 0.67 for the (λ, φ, β) = ($2.1\text{x}10^{-3}$, 2.0, 0.3) scenario. It is also interesting to note that based on discussions in the previous section regarding the majority of the error ratio being contributed by $IM < \eta_Z$ one would expect that if the power-model hazard is fit tangentially at a rate greater than that of $IM = \eta_Z$ the error would be smaller than that which occurs when the power-model hazard is fit at a rate less than $IM = \eta_Z$. Thus, it would be expected that the $IM_{2\lambda}$ fitting method is better than the $IM_{0.5\lambda}$ fitting method. Rows 6 and 7 of Table 1 illustrates that this assumption is not correct, in fact one would argue that based on Table 1 the $IM_{0.5\lambda}$ fitting method is better than the $IM_{2\lambda}$ fitting method. Figure 6c and Figure 6d show the deaggregation of the collapse hazard for the case of (λ, φ, β) = ($1.1\text{x}10^{-4}$, 4.0, 0.3). Figure 6c illustrates that using the low error ratio for the $IM_{0.5\lambda}$ fitting method is due to 'subtractive cancellation' as the analytical solution under-predicts the contribution around the region where the power-model is fit, and over-predicts the region where $IM >> \eta_Z$. Figure 6d illustrates that in this case fitting the power-model tangentially at a rate less than that of the $\lambda(\eta_Z)$ results in over approximation of the integral over the entire range of IM values.

For the regression fitting method we solve the least squares optimisation problem with various weighting functions:

$$\text{Minimise } R = \sum_{i=1}^{n} w_i \left[\ln(\lambda_i) - \ln(\lambda(IM_i))\right]^2 \tag{9}$$

where λ_i = data points of ground motion hazard curve; $\lambda(\text{IM}_i)$ = value of λ obtained from the power-model parametric equation (Equation (3)); and w_i = the weighting function for data point i. It would seem logical that the weights would be directly proportional to the range of IM values which contribute to the integrand. This will be a function of the distance between the data point (IM_i, λ_i) and $IM = \eta_Z$, as well as the aleatory uncertainty in the collapse fragility curve, β. The weight will therefore be related to the number of standard deviations of IM

points from η_Z. As the weight should reduce as the number of standard deviations increases then we use the inverse of the number of standard deviations for the weighting function:

$$w_i = \left(\frac{\beta_{RZ}}{\ln(IM_i) - \ln(\eta_Z)} \right)^{\gamma} \tag{10}$$

where γ is a parameter which controls the degradation of the weights as the number of standard deviations increases which is varied in the analysis to follow. The value γ = 0 would give a uniform weight to all data points. It is found that values of γ from 1-3.5 produce reasonable approximations to the integral. Figure 7a illustrates the hazard curves which are obtained for several different γ values by determining the parameters of Equation (3) via the solution of Equation (9). It can be seen that as the value of γ increases the power-model hazard curve approaches the tangent to the raw ground motion hazard data. Table 1 (base case (iv)) gives the error ratios when these parameters for the power-model are used. The tabulated values are also shown graphically in Figure 7b. It is evident that as before the error ratios generally increase as a function of dispersion, β, however, the error ratio is no longer directly proportional to the curvature which occurred in the tangent fit case (this is also true for several of the other non-tangent fits in Table 1) This is due to the method employed to compute the curvature (which uses only the second derivative of the hazard in log-log space at a single point), which was adequate when using a tangent-based fit, but does not appear adequate here. From Figure 7b it is also seen that no clear value of γ gives error ratios consistently close to 1.0, although out of all of the values of γ, one would probably suggest that γ = 2.0 yields the best results.

EPISTEMIC UNCERTAINTY IN COLLAPSE HAZARD

As previously mentioned, when epistemic uncertainties are considered in (either or both of) the collapse fragility curve and the ground motion hazard curve, it is possible to compute

the epistemic uncertainty in the collapse hazard, $\sigma_{\ln\lambda_C}$ (Equation (5)). Epistemic uncertainties arise in the collapse fragility due to finite sample uncertainty (estimating the parameters of the collapse fragility curve based on a finite number of points) and from analysis modelling uncertainty (assumptions on soil-structure-interaction, hysteresis models, 3-dimensional effects etc.), while epistemic uncertainty in the ground motion hazard is due to assumptions in Probabilistic Seismic Hazard Analysis (PSHA) (e.g. type and parameters for attenuation relations, magnitude recurrence relationships etc.). In the following paragraphs examples are given of the computation of the epistemic uncertainty (and the resulting distribution) in the collapse hazard using both the exact and closed-form solutions.

To compute the epistemic uncertainty in the 'exact' numerical solution, 5000 Monte-Carlo (MC) simulations were used (which was checked manually to verify it was sufficient for convergence of the non-parametric distribution). In the MC simulation the median *IM* causing collapse, η_z, and the ground motion hazard, $\lambda(IM)$, are assumed to be lognormal random variables as stated to obtain the closed-form solution for $\sigma_{\ln\lambda_C}$ (Equation (5)). Figure 8a illustrates the empirical CDF using epistemic uncertainties of (β_{UZ}, β_{UIM}) = (0.4, 0.3) which are typical epistemic uncertainties appearing in literature [8, 34]. As the actual ground motion hazard is used in the exact solution (as opposed to the power-model approximation) the distribution of the collapse hazard no longer has a lognormal distribution (which is the case for the closed-form solution). It is seen in Figure 8a that while a lognormal distribution (based on the sample median and standard deviation) is an adequate approximation over the central region of the distribution, its accuracy diminishes toward the tails of the distribution. It is also apparent that the magnitude of the epistemic dispersion, $\sigma_{\ln\lambda_C}$, is significant (a value of $\sigma_{\ln\lambda_C}$ =1.75 means that assuming a lognormal distribution, the 84th percentile collapse rate is 33 times the 16th percentile collapse rate, and that the 90th percentile is 3.85 times more than of the mean). This large epistemic dispersion is consistent with the closed-form solution, in

which the k^2 term amplifies the effect of the epistemic uncertainty in the collapse fragility curve, β_{UZ}.

It would seem intuitive that if the error ratio (E_{ratio}) in the closed-form solution for the expectation of λ_C is significant, then the error in $\sigma_{\ln\lambda_C}$ will also be significant. Of more importance however is: if the error λ_C from the closed-form solution is small, then will the error in $\sigma_{\ln\lambda_C}$ also be small? Possible reasons for significant error in $\sigma_{\ln\lambda_C}$ when E_{ratio} is small could be due to the aforementioned 'subtractive cancellation' in the expectation of the collapse hazard. Consider a single case using regression to fit the power-model (to the Christchurch hazard) with $\gamma = 2.0$, and using fragility and hazard parameters of $\beta_{RZ} = 0.3$, $\beta_{UZ} = 0.4$ and $\beta_{UIM} = 0.3$, $\varphi = 4.0$, respectively. These values are those used to obtain Figure 8a and from Table 1 give an error ratio of 1.0 for the expectation of the mean collapse rate. Using the regression approach with $\gamma = 2.0$ gives $k = 3.79$, and thus Equation (5) gives $\sigma_{\ln\lambda_C} = 1.54$. This is a 12% error compared to the actual value of 1.75 given in Figure 8a. Figure 8b illustrates the effect of the underestimation of the dispersion on the distribution of the collapse hazard. It is evident that the error in the dispersion primarily induces error in the collapse hazard for smaller levels of confidence. For example, the 12% error in the dispersion (Figure 8b) gives an error of 150% in predicting the median (with respect to epistemic uncertainties) value of the collapse hazard.

DISCUSSION

Numerous methods have been considered for the determination of the parameters of the power-model ground motion hazard. Although for each specific scenario it is possible to find a method for determining the parameters which gives a small error ratio, it has been rigorously shown that no method in general is adequate over the large range of likely values of the factors primarily influencing the error. From these results it is apparent that the

accuracy of the collapse hazard closed-form solution is very sensitive to the values of k and k_0 used (especially when the values of β and ϕ are significantly large). Hence if the closed-form solution is to be used then a great amount of care should be taken in selecting the values of these parameters. Based on the results of the parametric study it should be noted that there is unlikely to be any significant error when the closed-form solution is used to predict the annual rate of collapse for collapse-prone structures (i.e. those with an annual rate of collapse around $\lambda = 1x10^{-2}$). This is because Figure 3b illustrates that for frequent events, ϕ is typically less than 2, and Figure 5 shows that the error for this range of φ is small. Also, the error is strongly a function of the dispersion in the collapse fragility curve. This dependence on the dispersion further illustrates the need for advanced IM's which can accurately predict the effects of inelasticity and higher modes in complex structural behaviour [8, 26-28].

Another potential problem with the closed-form solution in its current form, as given in References [6-12], is that since the error is sensitive to the value of k used, in design environments either: (i) a large number of k values would have to be provided at different exceedance rates; or (ii) the raw hazard data would have to be provided, and designers should perform the necessary curve-fitting to obtain the value of k. It is likely, however, that the effort of the user to perform the power-model fit of the ground motion hazard (particularly if regression is used) is more than that required to directly numerically integrate Equation (1).

As a final remark, the results presented in this manuscript for the error between the closed-form solution for the annual rate of collapse and the direct numerical solution are also insightful toward the errors in the closed-form demand hazard solutions given in [7, 12, 29]. It is already acknowledged from previous work (e.g. [12]) that the simplifying assumptions necessary for arriving at the closed-form solution of the demand hazard could lead to significant error if the region in which the local approximations are made is distant from the region of major contribution to the integral (i.e. extrapolation from the region of parameter

fitting). This work has investigated the error in the closed-form solution for the collapse hazard where the above comments regarding extrapolation are not applicable. It has been shown that no method of determining the parameters of the power-model for the ground motion hazard, k_0, k, is in general, accurate over a range of ground motion hazard curves and collapse fragility curves which are likely to occur in practise. In addition to the error associated with the power-model representation of ground motion hazard, the demand hazard closed-form solution also assumes a power-model for the median demand-intensity relationship and constant logarithmic standard deviation. These additional two assumptions will introduce further error in the demand hazard [12] in addition to the assumptions in the ground motion hazard.

CONCLUSIONS

This study has investigated the error associated with the assumptions necessary to obtain the closed form for the annual frequency of structural collapse. The potential sources contributing to the error between the closed-form analytical solution and the exact solution for the annual rate of structural collapse were identified to be the curvature of the ground motion hazard and the dispersion in the collapse fragility curve, and the influence of these sources was been investigated via a parametric study. It was shown that the error in the closed-form solution is very sensitive to the log-log slope of the ground motion hazard curve, k, used, and while several fitting methods can be used to determine the value of k none are effective over the large range of likely values of parameters used. While the closed form analytical solution for the annual frequency of structural collapse is without doubt insightful, considering that the numerical evaluation of the annual rate of collapse is straightforward, the authors recommend that future performance-based design methods should not consider use of the closed-form solution a necessity.

ACKNOWLEDGEMENTS

Financial support of the first author by the New Zealand Tertiary Education Commission is greatly appreciated.

REFERENCES:

[1] FEMA-273. NEHRP guidelines for the seismic rehabilitation of buildings, Federal Emergency Management Agency, Washington, DC 1997.

[2] FEMA-302. NEHRP recommended provisions for seismic regulations for new buildings and other structures, Federal Emergency Management Agency, Washington, DC 1997.

[3] FEMA-350. Recommended seismic design criteria for new steel momentframe buildings, SAC Joint Venture, Washington, DC 2000.

[4] FEMA-356. Pre-standard and commentary for the seismic rehabilitation of building, Federal Emergency Management Agency, Washington, DC 2000.

[5] SEAOC. Vision 2000: A Framework for Performance-based Design, Structural Engineers Association of California, Sacramento 1995.

[6] Kennedy RP and Short SA. Basis for seismic provisions of Doe-Std-1020, Lawrence Livermore National Laboratory and Brookhaven National Laboratory 1994.

[7] Cornell CA. Reliability-based earthquake-resistant design—the future, in *Eleventh World Conference on Earthquake Engineering*, Acapulco, Mexico, 1996.

[8] Shome N and Cornell CA. Probabilistic seismic demand analysis of nonlinear structures, Stanford University, Stanford, CA, Report No. RMS-35, RMS Program, 1999.

[9] Cornell CA, Jalayer F, Hamburger RO, and Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines, *Journal of Structural Engineering,* vol. 128, pp. 526–533, 2002.

[10] Zareian F and Krawinkler H. Assessment of probability of collapse and design for collapse safety, *Earthquake Engineering and Structural Dynamics,* vol. 36, pp. 1901-1914, 2007.

[11] Mackie KR and Stojadinovic B. Performance-based seismic bridge design for damage and loss limit states, *Earthquake Engineering & Structural Dynamics,* vol. 36, pp. 1953-1971, 2007.

[12] Aslani H and Miranda E. Probability-based Seismic Response Analysis, *Engineering Structures,* vol. 27, pp. 1151-1163, 2005.

[13] Bradley BA, Dhakal RP, Cubrinovski M, Mander JB, and MacRae GA. Improved Seismic Hazard Model with Application to Probabilistic Seismic Demand Analysis, *Earthquake Engineering & Structural Dynamics,* vol. 36, pp. 2211-2225, 2007.

[14] Aslani H and Miranda E. Delivering improved information on seismic performance through loss deaggregation, in *Eighth National Conference on Earthquake Engineering*, San Francisco, CA, 2006, p. paper no. 1126.

[15] Bazzurro P and Cornell CA. Disaggregation of seismic hazard, *Bulletin of the Seismological Society of America,* vol. 89, pp. 501-520, 1999.

[16] Baker JW, Cornell CA, and Tothong P. Disaggregation of Seismic Drift Hazard, in *9th International Conference on Structural Safety and Reliability (ICOSSAR09)*, Rome, Italy, 2005.

[17] Ang AHS and Tang WH. *Probability Concepts in Engineering Planning and Design* vol. Volume I – Basic Principles: John Wiley & Sons, Inc., 1975.

[18] Jalayer F, Direct probabilistic seismic analysis: Implementing non-linear dynamic assessments, Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, Stanford, CA., 2003, 173 pp.

[19] Vamvatsikos D and Cornell CA. Incremental dynamic analysis, *Earthquake Engineering & Structural Dynamics,* vol. 31, pp. 491–514, 2002.

[20] Mander JB, Dhakal RP, Mashiko N, and Solberg KM. Incremental dynamic analysis applied to seismic financial risk assessment of bridges, *Engineering Structures,* vol. 29, pp. 2662-2672, 2007.

[21] Vamvatsikos D and Cornell CA. Applied incremental dynamic analysis, *Earthquake Spectra,* vol. 20, pp. 523–553, 2004.

[22] Ibarra LF and Krawinkler H. Global collapse of frame structures under seismic excitations, University of California at Berkeley, Berkeley, CA 2005.

[23] Stirling MW, McVerry GH, and Berryman KR. A new seismic hazard model for New Zealand, *Bulletin of the Seismological Society of America,* vol. 92, pp. 1878–1903, 2002.

[24] Baker JW and Cornell CA. A vector-valued ground motion intensity measure consisting of spectral acceleration and eplison, *Earthquake Engineering and Structural Dynamics,* vol. 34, pp. 1193-1217, 2005.

[25] Goulet CA, Haselton CB, Mitrani-Reiser J, Beck JL, Deierlein GG, Porter K, and Stewart JP. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses, *Earthquake Engineering & Structural Dynamics,* vol. 36, pp. 1973-1997, 2007.

[26] Luco N and Cornell CA. Seismic drift demands for two SMRF structures with brittle connections, in *Structural Engineering World Wide*: Elsevier: Oxford, England,, 1998.

[27] Tothong P and Luco N. Probabilistic Seismic Demand Analysis Using Advanced Ground Motion Intensity Measures, *Earthquake Engineering & Structural Dynamics,* vol. 36, pp. 1837-1860, 2007.

[28] Luco N and Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions, *Earthquake Spectra,* vol. 23, 2007.

[29] Kennedy RP, Cornell CA, Campbell RD, Kaplan S, and Perla HF. Probabilistic seismic safety study of an existing nuclear power plant, *Nuclear Engineering and Design,* vol. 59, pp. 315–338, 1980.

[30] Haselton CB, Assessing Collapse Safety of Modern Reinforced Concrete Moment Frame Buidlings., Ph.D. Thesis, Department of Civil and Environmental Engineering Stanford University, Stanford, CA, 2007,

Table 1: Error ratios for various ground motion hazard curve fitting methods.

Type	Fit factor /location	$\varphi = 2.0$[1]		$\varphi = 3.0$[1]		$\varphi = 4.0$[1]	
		$\beta = 0.3$	$\beta = 0.5$	$\beta = 0.3$	$\beta = 0.5$	$\beta = 0.3$	$\beta = 0.5$
Base-case	tangent	1.16	1.78	1.28	2.72	1.51	5.49
(i)	k=0.85kt	1.01	1.20	1.03	1.47	1.08	2.14
	k=0.75k_t	0.93	0.96	0.92	1.03	0.88	1.24
(ii)	$IM_{0.5\lambda}$	0.67	1.34	0.78	2.39	0.99	5.95
	$IM_{2\lambda}$	2.05	2.54	2.15	3.39	2.38	5.67
(iii)	$IM_{10\lambda}, \eta_Z$	0.95	1.03	0.967	1.25	1.02	1.86
	$IM_{5\lambda}, IM_{0.2\lambda}$	0.89	1.31	0.99	2.02	1.20	4.09
(iv)	$\gamma = 1$	0.52	0.71	0.65	0.73	0.47	0.46
	$\gamma = 2$	0.79	1.01	0.67	1.24	1.65	1.00
	$\gamma = 2.5$	0.90	1.48	1.15	2.32	0.84	1.67
	$\gamma = 3$	0.98	1.65	1.27	3.11	1.02	2.69
	$\gamma = 3.5$	1.02	1.74	1.37	3.90	1.17	3.81

[1] φ = 2.0, 3.0 and 4.0 correspond to the Christchurch ground motion hazard at approximately $\lambda(IM) = 2.1\text{x}10^{-3}$, $3.5\text{x}10^{-4}$ and $1.0\text{x}10^{-4}$, respectively.

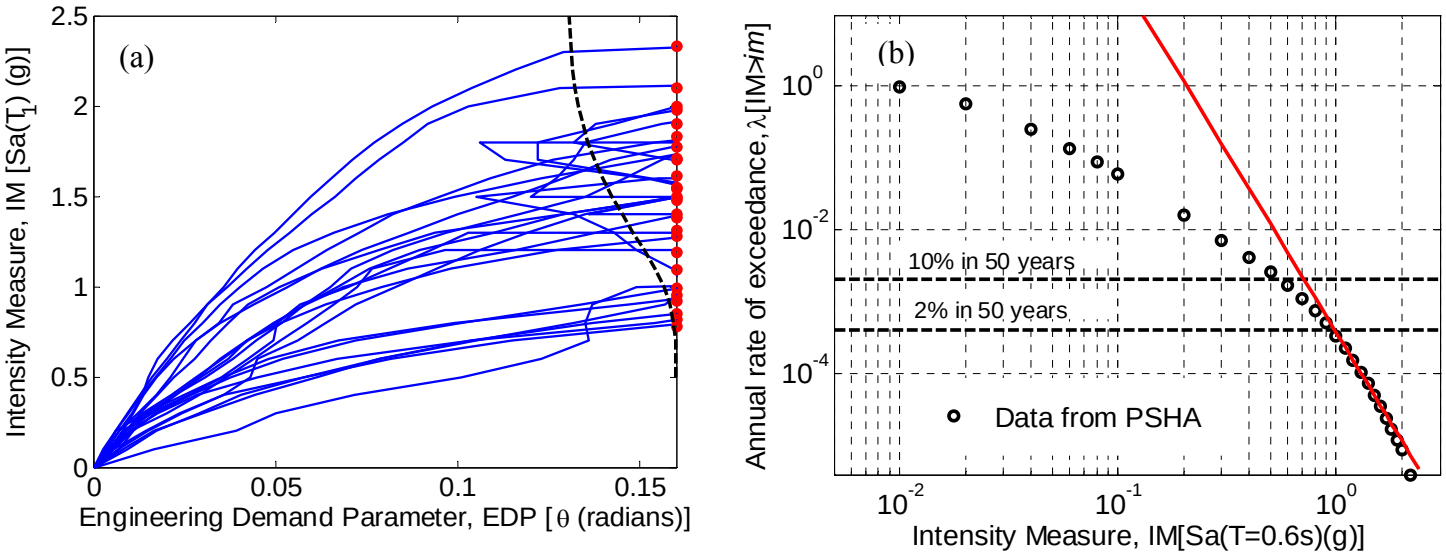

Figure 1: Comparisons between observational data and the parametric equations for the closed-form solution: (a) Seismic intensity-collapse relationship; and (b) Ground motion hazard

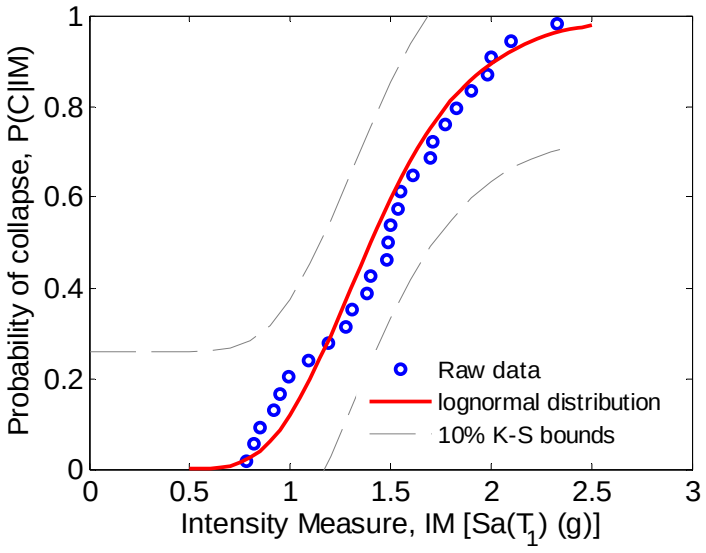

Figure 2: Collapse Fragility curve for the IDA curves in Figure 1a.

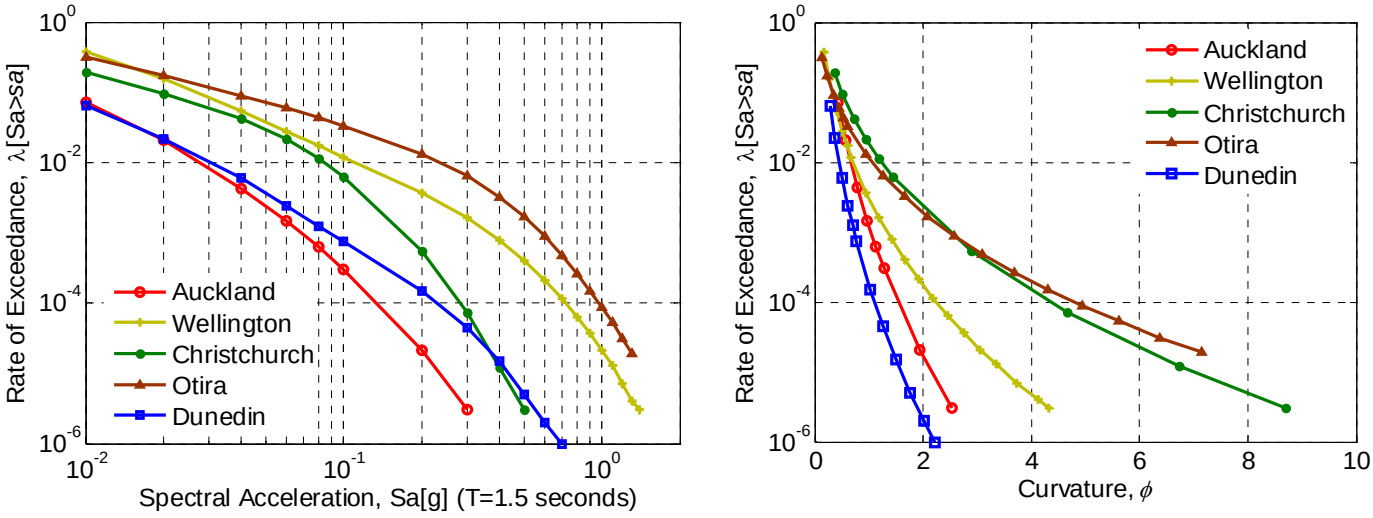

Figure 3: Hazard curves used in case study and curvature as a function of rate of exceedance: (a) Ground motion hazard curves; and (b) 'Curvature' of hazard curves

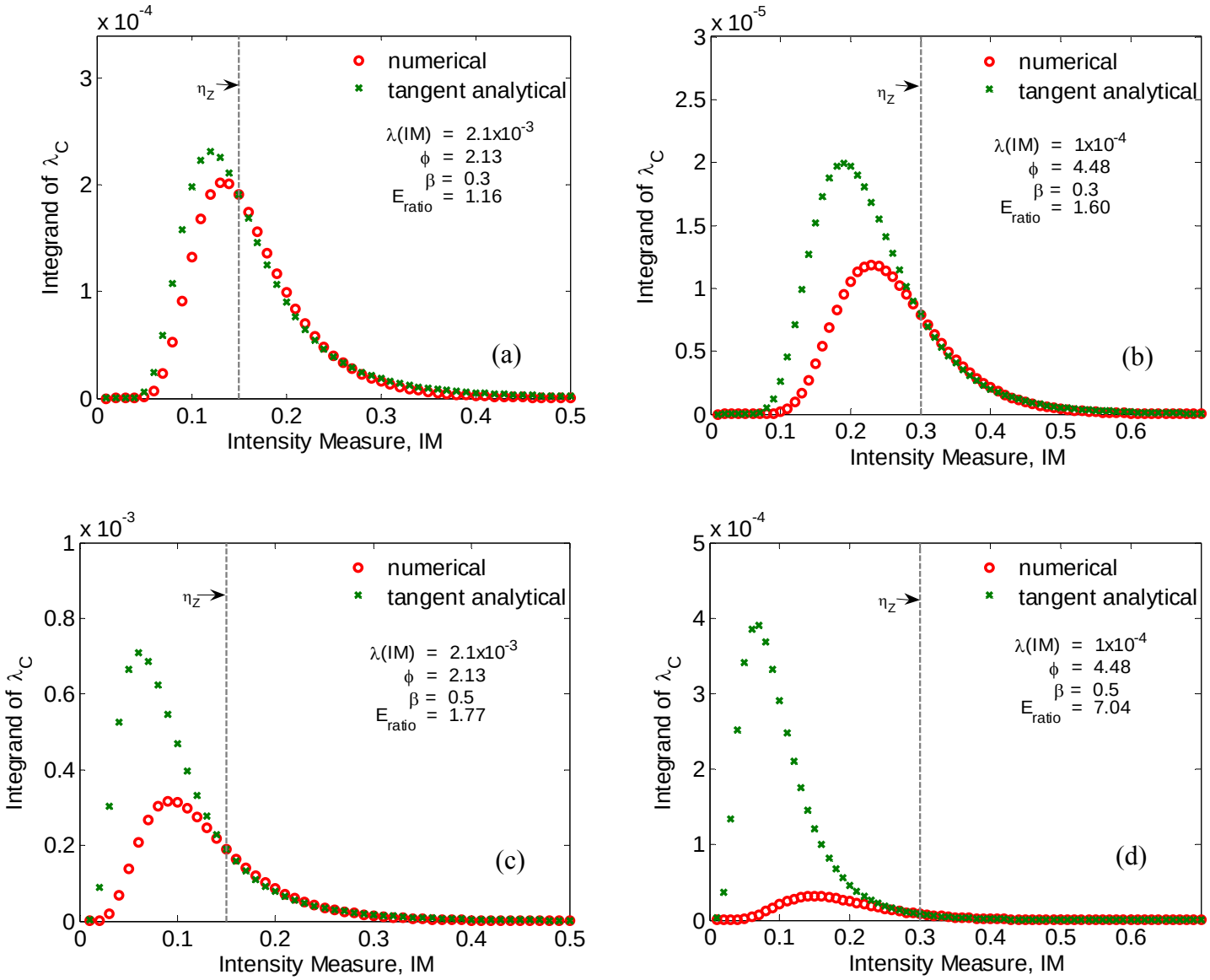

Figure 4: Deaggregation of Equation (1) for different values of φ and β.

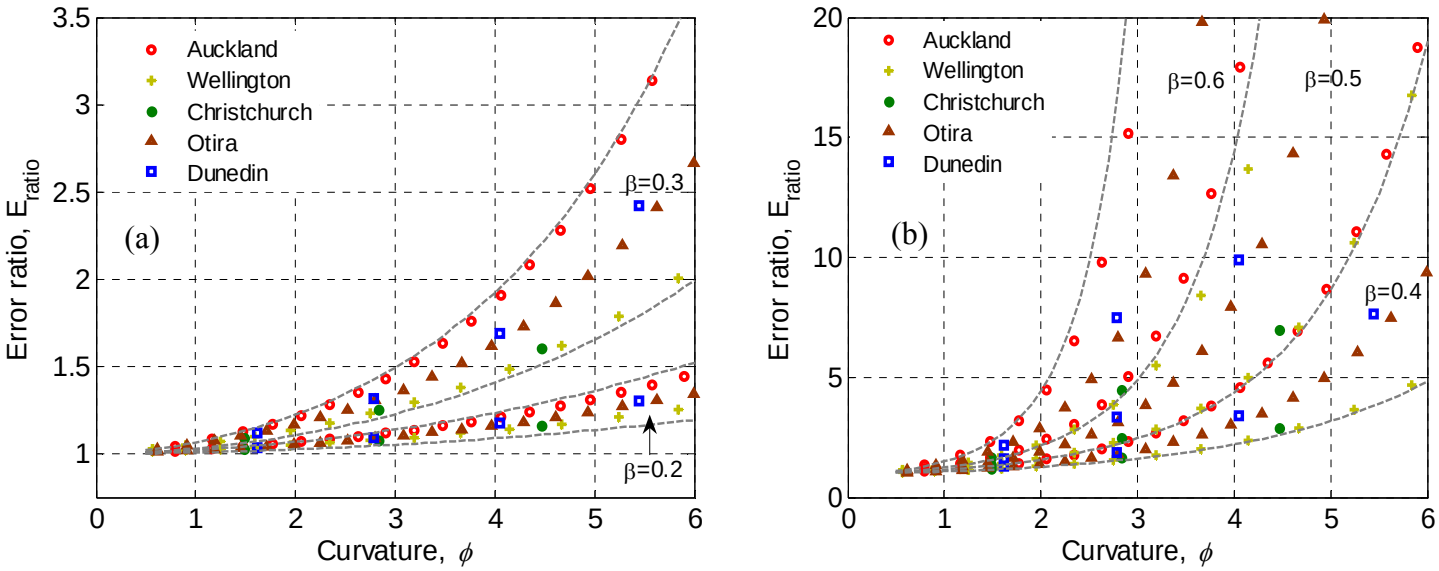

Figure 5: Error estimates for the tangent based approximation to the ground motion hazard curve for (a) β = 0.2 and 0.3; and (b) β = 0.4,0.5, and 0.6

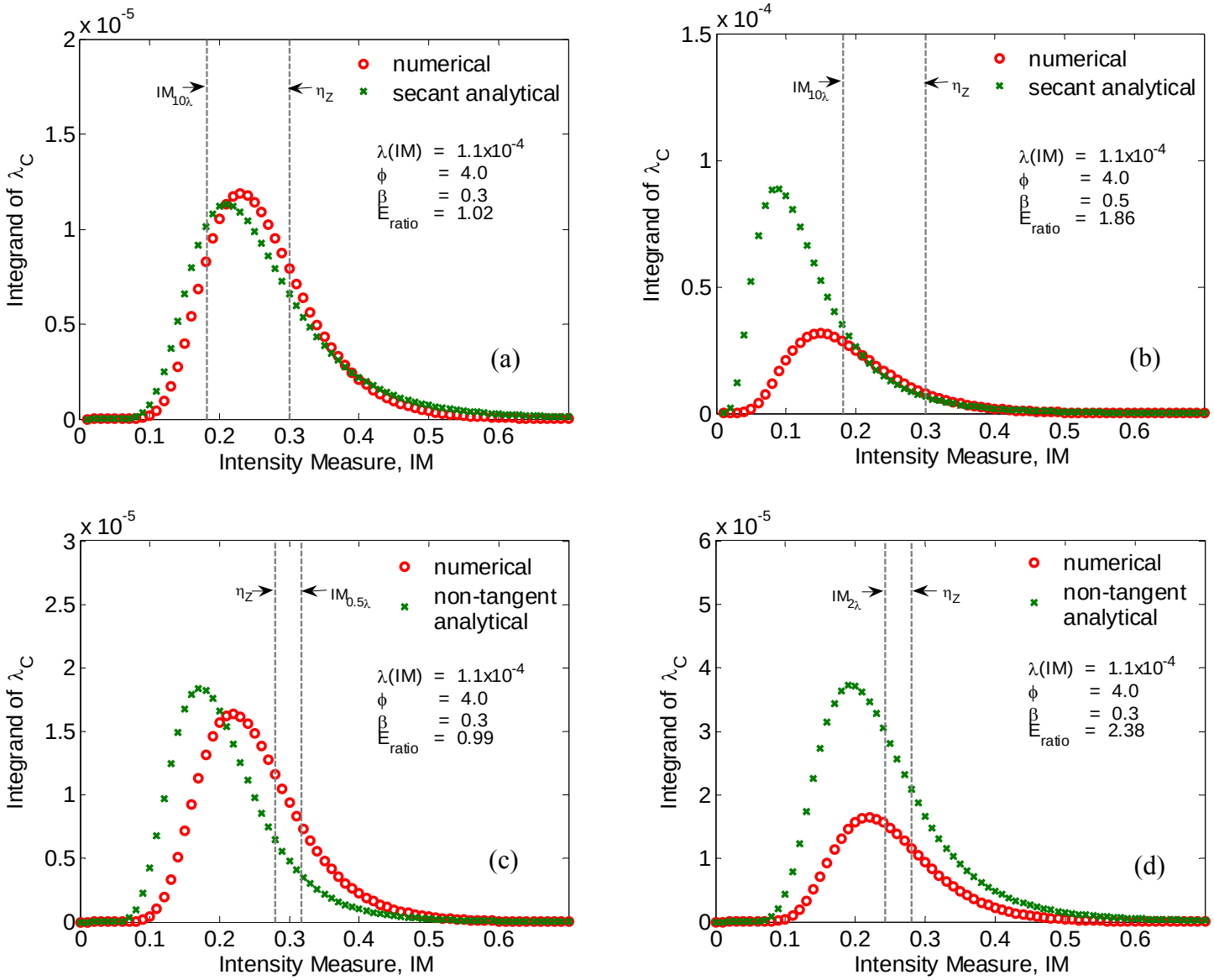

Figure 6: Deaggregation of Integral using secant-based power-model fit: (a) low error ratio due to sub-tractive cancellation;(b) non-occurrence of subtractive cancellation when β is increased; (c) low error ratio for $IM_{0.5\lambda}$ fit method; and (d) large error for $IM_{2\lambda}$ fit method.

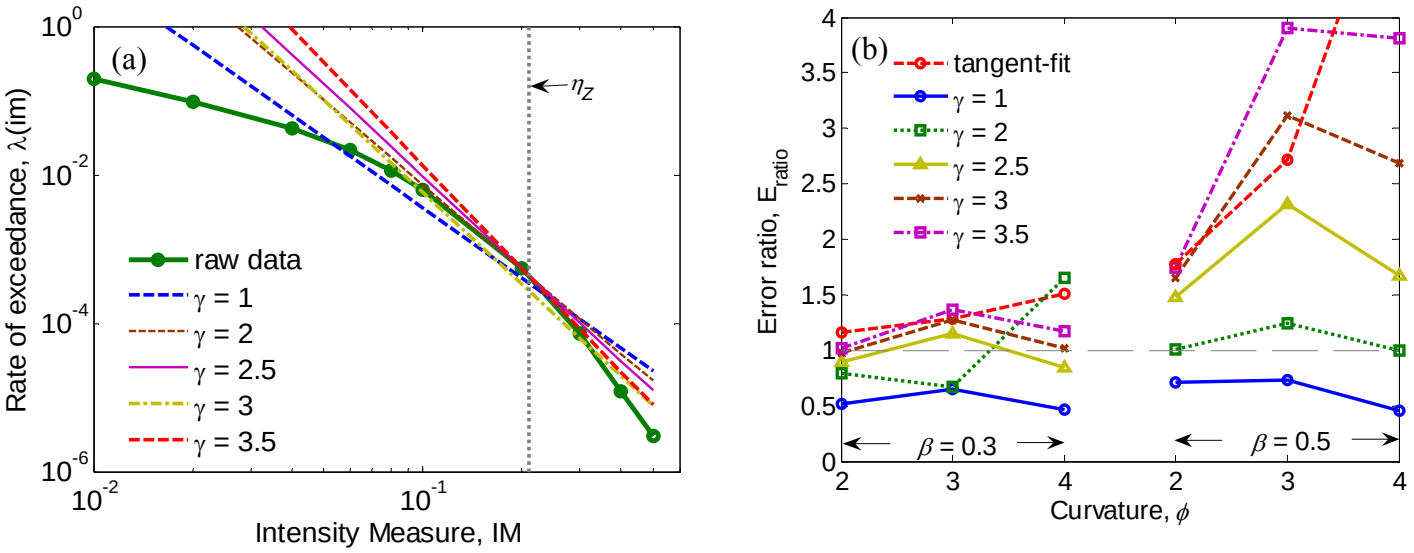

Figure 7: Illustration of using regression with various exponents for the weighting function: (a) effect on parameters, k, k_0 for β = 0.5 and φ = 3.0; and (b) error ratios.

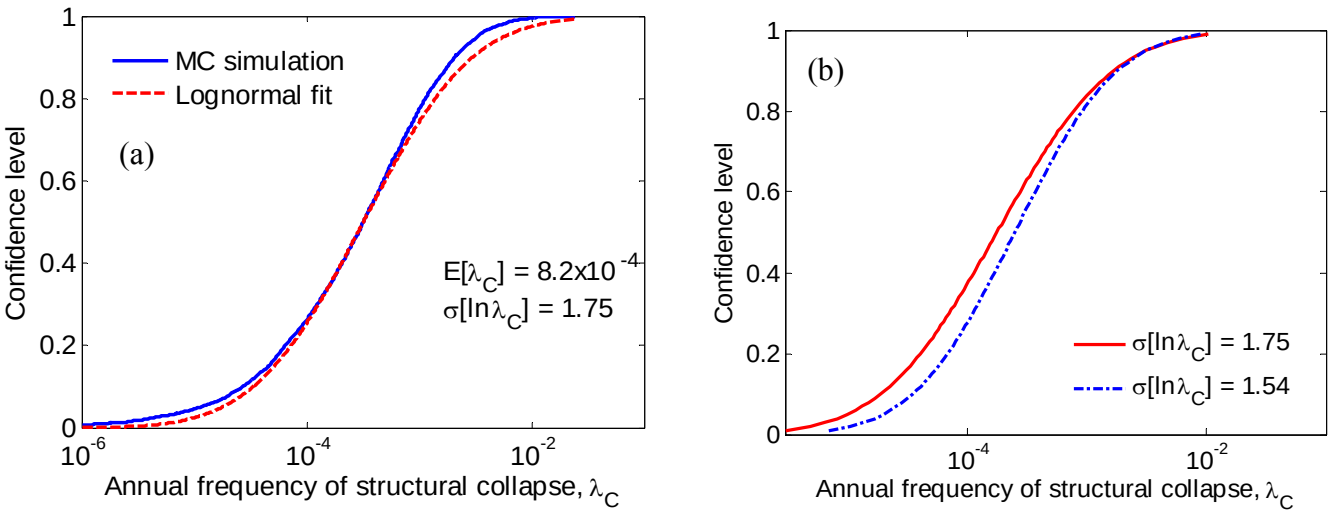

Figure 8: Error in collapse hazard distribution due to (a) lognormal parametric assumption of non-parametric distribution; and (b) underestimation of the epistemic dispersion.