Structural Safety 72 (2018) 74–83

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

Probability distribution of the seismic damage cost over the life cycle of structures

Mahesh D. Pandey *, J.A.M. van der Weide

Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

A R T I C L E I N F O

Article history:
Received 29 August 2016
Received in revised form 21 December 2017
Accepted 22 December 2017
Available online 30 December 2017

Keywords:
Stochastic point process
Poisson process
Seismic risk
Damage cost analysis
Probabilistic seismic hazard analysis
Life cycle analysis
Structural safety
Laplace transform

A B S T R A C T

In the life-cycle analysis, the total cost of damage caused by earthquakes is a significant but highly uncertain component. In the current literature, the seismic risk analysis is largely limited to the evaluation of the average cost of damage, which is not informative about the full extent of variability in the cost. The paper presents a systematic development of the stochastic modeling of seismic risk analysis problem and reformulates the damage cost analysis as a superposition of compound Poisson processes. An explicit analytical solution for the distribution of damage cost is derived in form of a recursive equation. The proposed approach extends the capability of the existing framework of seismic risk analysis, which can be used to optimize initial design and retrofitting of structures.

1. Introduction

1.1. Background

With the advent of probabilistic modeling of structural safety problems in 1960s and 70s, the risk-based calibration and optimization of design codes became an important area of research in structural engineering. Comprehensive studies by Whitman et al. [1] and Whitman and Cornell [2] laid the foundation of the seismic risk analysis and design of structures that are likely to face multiple seismic events during the service life. This framework, which still serves as a basis for modern code development [3], embodies the following key ideas:

- The two most basic goals of seismic design were recognized as minimizing the likelihood of costly repairs and avoiding the loss of life [4, p. 18].
- Stochastic modeling of earthquake occurrences as the homogeneous Poisson process accompanied by a random variable representing the ground motion intensity [5].
- The probability distribution of annual maximum of the ground motion intensity as a basis to evaluate structural reliability. This is to meet the objective of life safety.
- The expected cost of losses caused by seismic events as a measure of seismic risk, which in turn relates to the economic efficiency of a design.
- Optimization of seismic resistance by balancing the initial cost premium for improving the resistance against the future expected cost of damage. This guiding principle serves as a basis for code development.

In recent years, the move towards the performance-based design has prompted the development of more and more refined models for life cycle cost analysis in which the seismic damage cost is an important but fairly uncertain element. Note that he term *damage cost* includes all the losses that could incur due to loss of services, damage to contents and cost of repairing and restoring the damaged structure.

In the current literature, the life cycle analysis is almost exclusively focussed on the evaluation of expected (or average) cost of seismic damage. The expected cost is not informative about the extent of losses, given that a large variance is associated with the damage cost [6]. It means that an exclusive reliance on the expected cost for optimizing a design would not yield a realistic result due to large variability potentially associated with an outcome. Instead, an upper percentile of the cost would be a more meaningful bound of the risk, such as the Value at Risk (VaR) used in the financial literature. In spite of this limitation of the expected cost measure, the evaluation of probability distribution of damage

* Corresponding author.
E-mail address: mdpandey@uwaterloo.ca (M.D. Pandey).

https://doi.org/10.1016/j.strusafe.2017.12.007
0167-4730/© 2017 The Authors. Published by Elsevier Ltd.

Nomenclature

SMRF	steel moment resisting frame	PDF	probability density function
DCPP	Diablo Canyon power plant	PMF	probability mass function
PSHA	probabilistic seismic hazard analysis	F	cdf function
RV	random variable	*iid*	independent and identically distributed
CDF	cumulative distribution function	HPP	homogeneous Poisson process
POE	probability of exceedance	δ	unit cost of damage

cost has not been tackled in the seismic literature. It could well be a result of a presumption that the simulation-based method is the only way to determine the cost distribution, and that the computational burden associated with simulation makes it impractical tool foe seismic risk analysis.

Thus, the primary goal of this study is to present a clear exposition of the stochastic modeling of seismic risk analysis that leads to an analytical expression for the probability distribution of total cost of seismic damage. This distribution can be used to evaluate a probabilistic bound on risk, which could serve as a basis to optimize design and retrofitting options for structures.

The approach taken in this paper is to draw a parallel between the seismic risk analysis and the theory of stochastic renewal process. This understanding provides new interpretations and insights which are necessary to derive the full distribution of the damage cost.

1.2. A motivating example

The analytical formulations presented in this paper are aimed to analyze practical examples like the seismic damage cost for a 20-storey steel moment resisting frame (SMRF) building, as shown in Fig. 1. The building was designed as a standard office building sitting on stiff soil. It has a fundamental period of 4.0 s. Other structural details can be found elsewhere [7].

The building is (hypothetically) situated at a site of the Diablo Canyon power plant (DCPP) in California. This site is in the proximity of the Hosgri, Los Osos, San Luis Range, and Shoreline faults, as

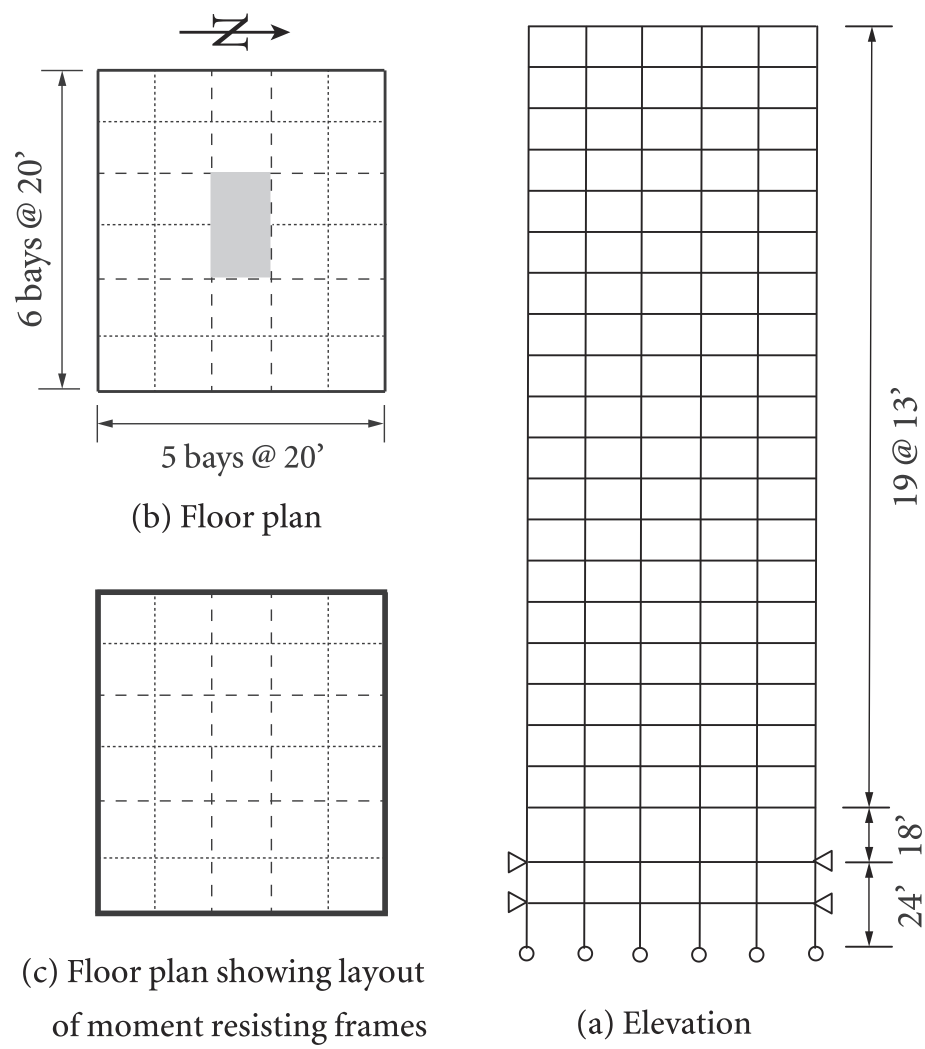

Fig. 1. Example building for seismic risk analysis.

shown in Fig. 2. It is vulnerable to seismic damage by 13 potential sources, the parameters of which are complied by Wang [8] from a number of references. The ground-motion prediction equations given by Abrahamson and Silva [9] were used in probabilistic seismic hazard analysis (PSHA) [8]. The ground motion hazard is quantified in terms of the spectral acceleration at the fundamental period of 4 s. The seismic hazard curve resulting from all seismic sources is shown in Fig. 3.

The objective of the analysis is to determine the probability distribution of the damage cost over a 50 year service life of the building. A complete solution of this problem is presented in Section 7 of the paper.

1.3. Technical problem definition

The stochastic model of a recurring hazard like earthquake includes two random variables, the time (T) between events and the magnitude of hazard (X). Moreover, the cost of damage resulting from an event, (C), being a function of X, is also taken as a random variable. Thus, a recurring hazard can be modelled as a sequence of random vectors $(T_i, X_i),\ i = 1, 2, \ldots$, which is technically referred to as a *marked point process* (see Fig. 4). The number of events, $N(t)$, in a time interval $(0, t]$ is also a stochastic variable, since it is a function of T_i.

This stochastic model is intended to derive the two important quantities. The first is the maximum hazard magnitude in a time interval $(0, t]$:

$$X_{max}(t) = \max(X_1, \ldots, X_{N(t)}), \quad (\text{if } N(t) \geqslant 1) \tag{1}$$

The probability distribution of $X_{max}(t)$ is required for structural reliability analysis.

The other quantity of interest is the cumulative or total damage cost resulting from all possible hazardous events in $(0, t]$, i.e.,

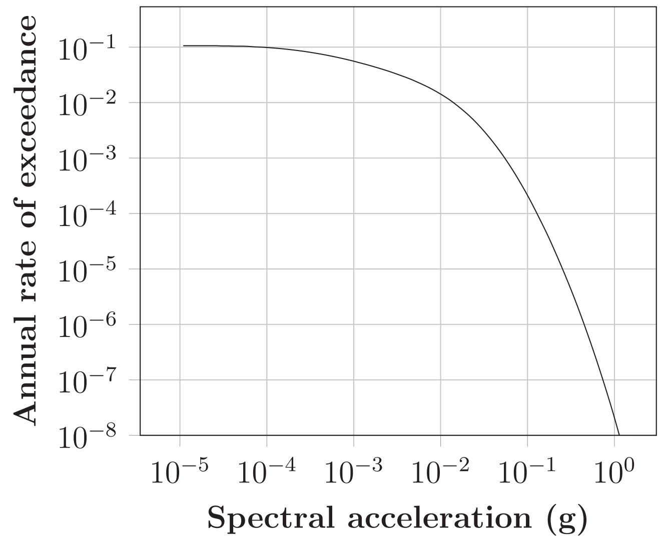

Fig. 3. Seismic hazard curve at the DCPP site (spectral acceleration at 4 s period of the structure).

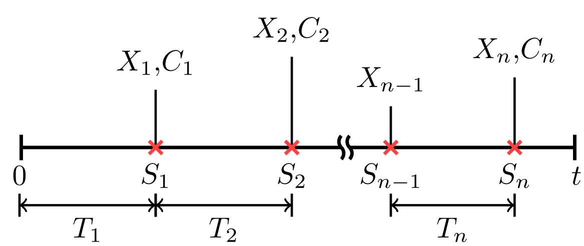

Fig. 4. Stochastic modeling of a recurring hazard.

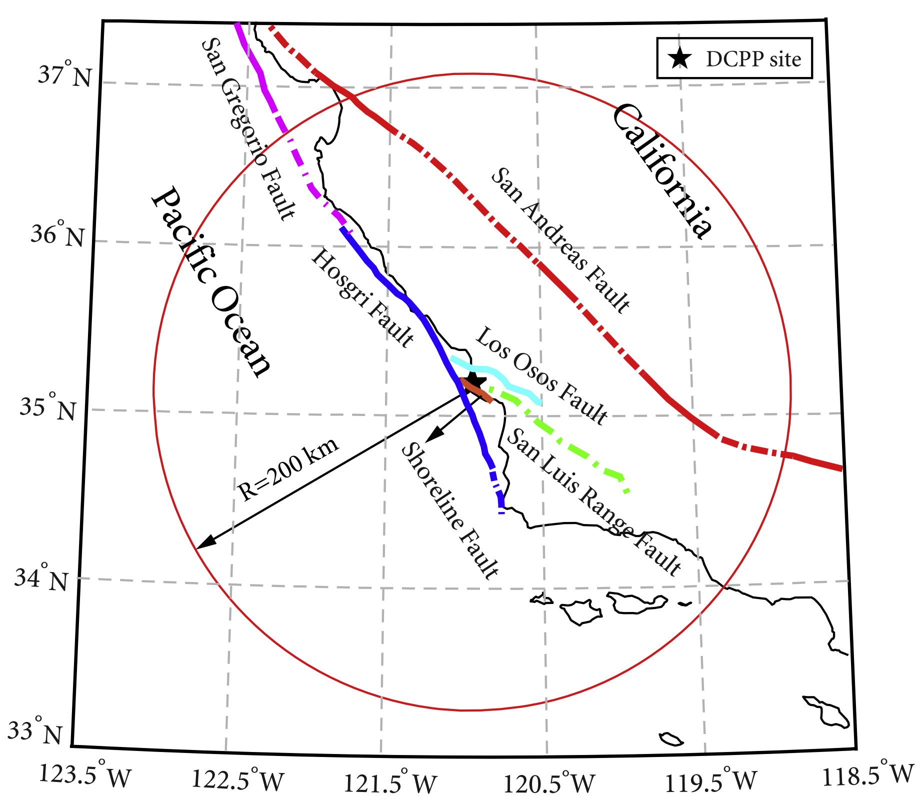

Fig. 2. Seismic sources for the DCPP site.

$$Y(t) = \sum_{i=1}^{N(t)} C_i \tag{2}$$

The expected value of total damage cost, $\mathbb{E}[Y(t)]$, is commonly known as "risk".

In the seismic risk analysis, $N(t)$ is modelled as a homogeneous Poisson process for which the distribution of $X_{max}(t)$ and the risk, $\mathbb{E}[Y(t)]$, are available in explicit analytical forms.

The paper shows that the distribution of $Y(t)$ can as well be derived in analytical terms, without the recourse to a simulation method. In addition, the paper clarifies several aspects of the current practice of seismic risk analysis.

Since a realistic analysis involves multiple seismic sources, the paper shows that the effect of multiple sources can be reduced to an equivalent single compound process via a superposition principle. This is a significant innovation of this study.

1.4. Background literature

Although there is a substantial body of literature on seismic risk and life cycle cost analysis of different types of structures, this Section refers to few key studies that are relevant to fundamental developments of the probabilistic methodology.

The maximization of expected utility as a basis to optimize the structural design was formulated by Esteva [10], who proposed expected discounted cost as a measure of the utility. The next comprehensive study on this topic was presented by Rosenblueth [11], who introduced a more general stochastic renewal process for estimating the expected discounted losses caused by stochastic hazards. The problem of optimum design of structures under dead, live and seismic loads was analyzed by Rosenblueth [12]. These studies firmly established the expected cost as a basis for optimization that is still continued to be used in the literature [13]. Application of the expected loss analysis to performance based design was presented by Krawinkler et al. [14]. In a series of papers, Rackwitz and his coworkers applied the renewal process model to a more general class of problems in which the effects of degradation, maintenance and societal benefits were included in life cycle cost analysis [15–17]. They presented formulations of increasing refinements for the evaluation of expected discounted cost over an infinite time horizon.

A few exceptions to the universal use of expected cost are seen in the work of Goda and Hong [18], who applied the Monte Carlo simulation method to evaluate the mean, standard deviation and probability distribution of the cost. Porter et al. [19] evaluated the variance of discounted damage cost. The expected utility analyses of different design options and risk mitigation programs were presented by mitigation programs were presented by Cha and Ellingwood [20].

There has been a great deal of theoretical and experimental research to understand the damaging effects of earthquakes and their relation to various types of losses and costs of repair. This has led to the development of more realistic models for the damage cost analysis [21,22]. The relation of seismic damage to economic losses was explored by Martins et al. [23]. Life-cycle analysis of steel buildings exposed to wind and seismic hazards was presented by Mahmaoud and Cheng [24], which was based on the expected cost formulation of Kang and Wen [25]. A study by Gosh and Padgett [26] considered the effect of aging and reported that seismic losses are significantly increased as a result of deterioration of structures. In addition to risk analysis, some researchers are focussing on the resilience in terms of recovery of operation followed by an earthquake [27].

Based on the review, it is concluded that the evaluation of the probability distribution of damage cost has not yet been analyzed in a comprehensive manner the seismic literature.

1.5. Organization

The paper is organized as follows. Section 2 presents the basic terminology and concepts of the stochastic process models, such as marked and compound processes. Section 3 shows that the homogeneous Poisson process is a special case of a renewal process for which several analytical results can be derived. Section 4 presents the main result of this paper by deriving the probability distribution of a compound Poisson process. Section 5 is a key part of this paper, as it connects the analytical results presented in Sections 2–4 to the evaluation of seismic risk in a practical setting, which considers multiple seismic sources affecting the site and the structure. This Section presents a novel approach to analyse this problem in a single step via the superposition properties of the compound Poisson process. Section 6 summarizes the application of analytical results presented in the paper to seismic risk analysis. Section 7 presents the solution of a practical example of risk analysis, which was introduced in Section 1.2. The last Section summarizes the key findings of this study.

2. Stochastic point processes: key concepts

In this Section, key concepts of the theory of stochastic point processes are reviewed that serve as a foundation of the seismic risk analysis.

2.1. Basic terminology

A random variable (RV) is denoted by a capital roman letter, such as X, and its cumulative distribution function (CDF) is denoted as $F_X(x) = \mathbb{P}[X \leqslant x]$. The complementary CDF or probability of exceedance (POE) function is defined as $\bar{F}_X(x) = 1 - F_X(x)$. The probability density function (PDF) of a continuous RV is denoted as $f_X(x)$ and the same notation is adopted for the probability mass function (PMF) of a discrete RV.

The sum of n independent and identically distributed (*iid*) random variables, $X_1, \ldots, X_n$, is denoted as, $S_n = X_1 + X_2 + \cdots + X_n$. The probability distribution of this sum is given by an n-fold convolution:

$$F_{S_i}(x) = \mathbb{P}[X_1 + X_2 \cdots + X_n \leqslant x] = F_X^{(n)}(x), \tag{3}$$

which can be recursively evaluated as [28]:

$$F_X^{(n)}(x) = \int_0^x F_X^{(n-1)}(x-y)\, \mathrm{d}F_X(y), \quad (n \geqslant 1) \tag{4}$$

Note that $F_X^{(1)}(x) = F_X(x)$, $F_X^{(0)}(x) = 1$, and $\mathrm{d}F_X(x) = f_X(x)\mathrm{d}x$ when the probability density of X exists. For $n \gg 1$, a convolution, $F_X^{(n)}(x)$, is not easy to compute due to numerical difficulties associated with the evaluation of a higher order recursive integral.

2.2. Point process, counting process and renewal process

Mathematically, a (simple) point process is a random and strictly increasing sequence, $S_0 = 0 < S_1 < S_2 < \cdots$, of positive RVs without a finite limit point i.e. lim $S_n \to \infty$ as $n \to \infty$. A point process can be equivalently represented by a sequence of *random* inter-occurrence time (Fig. 5) $T_1, T_2, \ldots$, with $T_n = S_n - S_{n-1}$. The arrival time of an ith event, S_i, can be written as a partial sum, $S_i = T_1 + T_2 \cdots + T_i$.

The number of events in the time interval $(0, t]$, denoted as $N(t), t > 0$, is referred to as the *counting process* associated with the partial sums S_i, $i \geqslant 1$. It is defined in terms of an indicator function as, $N(t) = \sum_{i=1}^{\infty} \mathbf{1}_{\{S_i \leqslant t\}}$. Note that an indicator function is defined as, $\mathbf{1}_A = 1$ if A is true, otherwise $\mathbf{1}_A = 0$. Since the events $\{N(t) = i\}$

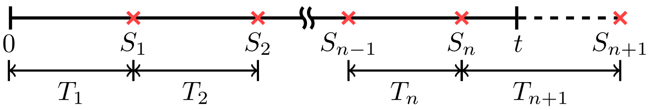

Fig. 5. A schematic of the renewal process.

and $\{S_i \leqslant t < S_{i+1}\}$ are identical, the marginal PMF of $N(t)$ can be written as

$$f_{N(t)}(i) = \mathbb{P}[N(t) = i] = \mathbb{P}[S_i \leqslant t < S_{i+1}] = F_{S_i}(t) - F_{S_{i+1}}(t), \quad i = 0, 1, \ldots \tag{5}$$

A point process is called an ordinary *renewal process* if the inter-occurrence times $T_1, T_2, \ldots$,form a sequence of non-negative, *iid* RVs with a distribution $F_T(t)$. The word "renewal" implies that the process is reset after each occurrence of the event of interest.

The renewal function, $\Lambda(t)$, is a key summary measure of the renewal process, and it is defined as the expected number of renewals in a time interval $(0, t]$. The renewal function is evaluated via an integral equation [28]:

$$\Lambda(t) = F_T(t) + \int_0^t \Lambda(t-y)\, \mathrm{d}F_T(y).$$

The renewal rate is the expected number of renewals per unit time, i.e., $\lambda(t) = \mathrm{d}\Lambda(t)/\mathrm{d}t$.

2.3. Marked renewal processes

A marked renewal process, as shown in Fig. 6, is defined as a sequence of *iid* random vectors (T_i, X_i), $i = 1, 2, \ldots$ [29, Section 6.2, p. 321]. The maximum hazard magnitude, $X_{max}(t)$, in a interval, $(0, t]$ is defined as

$$X_{max}(t) = \begin{cases} \max(X_1, \ldots, X_{N(t)}), & \text{if } N(t) \geqslant 1, \\ 0 & \text{if } N(t) = 0 \end{cases} \tag{6}$$

Using the total probability theorem, the distribution of $X_{max}(t)$ can be written as

$$F_{X_{max}}(x; t) = \sum_{n=1}^{\infty} \mathbb{P}[\max(X_1, \ldots, X_n) \leqslant x, N(t) = n] + \mathbb{P}[N(t) = 0] \tag{7}$$

2.4. Compound renewal processes

The cumulative effect of a marked renewal process in $(0, t]$ is called the *compound renewal process*, $Y(t)$, as defined by (2). Its distribution can be written using the total probability law for any $y > 0$ as:

$$\begin{aligned} \mathbb{P}[Y(t) \leqslant y] &= \mathbb{P}\left[S^X_{N(t)} \leqslant y\right] \\ &= \mathbb{P}[N(t) = 0] + \sum_{n=1}^{\infty} \mathbb{P}\left[S^X_n \leqslant y, N(t) = n\right] \end{aligned} \tag{8}$$

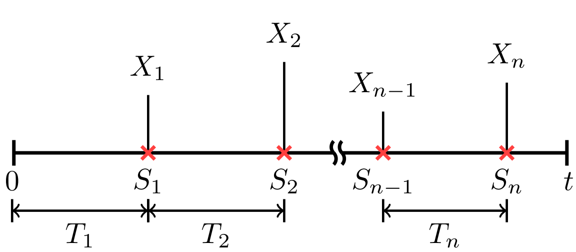

Fig. 6. An example of a marked renewal process.

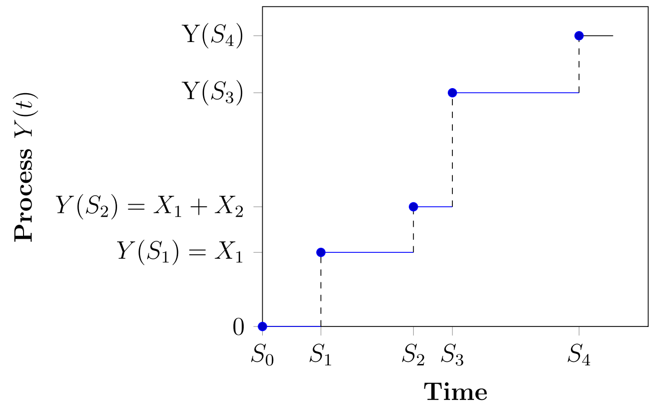

Fig. 7. A sample path of a compound renewal process.

where $S^X_n = \sum_{i=1}^n X_i, n \geqslant 1$. A sample path of a compound process is shown in Fig. 7. If T_i and X_i are independent, the following expression can be obtained:

$$F_{Y(t)}(y) == f_{N(t)}(0) + \sum_{n=1}^{\infty} F^{(n)}_X(y) f_{N(t)}(n) \tag{9}$$

The above expression is not suitable to compute the distribution of $Y(t)$, as it involves an infinite sum of n-fold convolutions.

2.5. Remarks

The distributions associated with renewal processes, as described in this Section, do not admit any analytical solutions when T and X follow general types of probability distributions. For example, the renewal function involves an infinite sum of convolutions of T. The compound process, in addition includes an infinite sum of convolutions of X, which are not easy to compute. Nevertheless, the Poisson process is a special case for which several analytical solutions can be obtained, as shown in the later parts of the paper.

3. Models based on the homogeneous Poisson process (HPP)

3.1. Renewal process

The homogeneous Poisson process (HPP) is the simplest and most widely used renewal process in which the time between events is an exponentially distributed RV with the distribution, $F_T(t) = 1 - e^{-\lambda t}$. The distribution of $N(t)$ is explicitly given by the Poisson PMF,

$$\mathbb{P}[N(t) = n] = f_{N(t)}(n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, \quad n = 0, 1, 2, \ldots \tag{10}$$

This can be written in a recursive form as

$$\begin{cases} f_{N(t)}(0) = e^{-\lambda t} \\ f_{N(t)}(n) = \frac{\lambda t}{n} f_{N(t)}(n-1), & n \geqslant 1 \end{cases} \tag{11}$$

All statistical moments of an HPP can be conveniently derived from the method of the Laplace transform. The renewal function is given as, $\mathbb{E}[N(t)] = \Lambda(t) = \lambda t$ with a constant renewal rate, λ.

3.1.1. The decomposition property

Suppose events are generated by an HPP, $(N(t); \lambda)$, and on arrival of each event, there is a probability p of the event being recorded. In this manner, the original process is decomposed into two sets, recorded and not recorded events.

The decomposition property means that random selections from a Poisson process of rate λ are equivalent to a Poisson process with rate λp [29, Chapter V].

3.2. Marked Poisson process

The seismic risk analysis is based on this stochastic representation of a single seismic source, which has potential to generate seismic waves in a given time interval. The seismic hazard curve, $H(x)$, is defined as the rate of occurrence of earthquakes exceeding certain magnitude, x. Based on the decomposition property of the Poisson process, it can be written as

$$H(x) = \lambda\, \overline{F}_X(x) = \lambda\, (1 - F_X(x)) \tag{12}$$

Substitution of the basic properties of HPP into Eq. (7) leads to a rather well known result for the distribution of maximum value:

$$F_{X_{max}}(x;t) = \sum_{n=1}^{\infty} (F_X(x))^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} + e^{-\lambda t} = e^{-\lambda t \overline{F}_X(x)}. \tag{13}$$

This distribution corresponds to annual maximum value when $t = 1$ [5, Eq. (22)]:

$$F_{X_{max}}(x;1) = e^{-\lambda\, \overline{F}_X(x)} = e^{-H(x)}, \tag{14}$$

which is directly related to the hazard curve defined by Eq. (12).

It is important to point out that this distribution has an atom at 0, i.e., $F_{X_{max}}(x = 0; t = 1) = e^{-\lambda}$. Using a first order approximation, $e^a \approx 1 + a$ when $a \ll 1$, the distribution of annual maximum value can be approximated as $F_{X_{max}}(x;1) \approx 1 - \lambda\, \overline{F}_X(x)$, such that its POE function can be approximated by the hazard curve as

$$\overline{F}_{X_{max}}(x) \approx \lambda \overline{F}_X(x), \quad \text{or } \overline{F}_{X_{max}}(x) \approx H(x) \tag{15}$$

3.2.1. Remarks

The approximation of the annual maximum distribution by the hazard curve is quite commonly used in the seismic risk analysis. This concept is so ingrained in the practice that the hazard curve, $H(x)$, has taken the place of the probability distribution, and $dH(x)/dx$ is treated as probability density, which is technically incorrect.

It is important to clarify that $F_{X_{max}}(x;1)$ is a probability, whereas $H(x)-$ a rate, which are two different quantities. The approximate equivalence between them holds only in a numerical sense, not in any conceptual sense. This misconception can be a source of error in the analysis of total damage cost.

3.3. Compound Poisson process: computation of risk

Using Eq. (9), the distribution of the compound Poisson process can be written as

$$F_{Y(t)}(y) == e^{-\lambda t} + \sum_{n=1}^{\infty} F_X^{(n)}(y)\, \frac{(\lambda t)^n}{n!} e^{-\lambda t} \tag{16}$$

This expression is not suitable to compute the distribution, as it involves an infinite sum of n-fold convolutions. However, the moments of $Y(t)$ can be easily derived from the method of Laplace transform [28]. Especially, the expected value has a fairly simple expression, $\mathbb{E}[Y(t)] = \lambda t\, \mathbb{E}[X]$. This simplicity is one of the main reasons for popular use of the HPP model in the probabilistic life-cycle analysis.

4. Probability distribution of the compound Poisson process

As discussed earlier, an analytical solution for the distribution of a general compound process can not be derived. However, the compound Poisson process admits a solution based on the properties of the Poisson distribution and the discrete convolution. This Section shows that a simple recursive equation can be derived to compute the distribution of $Y(t)$. This recursion scheme works when X is a discrete RV or it is a suitably discretized representation of a continuous RV [30].

Suppose X represents the monetary loss due to a single event and it has a PMF

$$P(X = i\delta) = f_X(i\delta), \quad i = 0, 1, 2, \ldots,$$

where δ is the unit cost, say a thousand dollar. Based on the discussion of Section 2.1, the PMF of the partial sum, $S_n = \sum_{i=1}^{n} X_i$, $n \geqslant 1$, can be written as

$$f_X^{(n)}(i\delta) = \sum_{j=0}^{i} f_X^{(n-1)}((i-j)\delta) f_X(j\delta) \tag{17}$$

Note the convention used here, $f_X^{(0)}(i\delta) = 0$ if $i \neq 0$ and $f_X^{(0)}(0) = 1$. The PMF, $f_{Y(t)}(i\delta),\ i \geqslant 0$, can be written using the total probability law as

$$f_{Y(t)}(i\delta) = \sum_{n=0}^{\infty} \mathbb{P}[S_n = i\delta] f_{N(t)}(n) = \sum_{n=0}^{\infty} f_X^{(n)}(i\delta) f_{N(t)}(n) \quad (\text{for } i \geqslant 0) \tag{18}$$

For a non-zero value of $Y(t) = i\delta$, the PMF can be written as

$$f_{Y(t)}(i\delta) = \sum_{n=1}^{\infty} f_X^{(n)}(i\delta) f_{N(t)}(n) \quad (\text{for } i \geq 1) \tag{19}$$

Substituting the recursive relation for $f_{N(t)}(n)$ from Eq. (11) leads to

$$f_{Y(t)}(i\delta) = \sum_{n=1}^{\infty} \frac{\lambda t}{n} f_X^{(n)}(i\delta) f_{N(t)}(n-1) \tag{20}$$

For a discrete convolution, the following relation is useful [28]:

$$\frac{1}{n} f_X^{(n)}(i\delta) = \sum_{j=1}^{i} \frac{j}{i} f_X(j\delta) f_X^{(n-1)}((i-j)\delta), \quad (\text{for } i, n \geqslant 1). \tag{21}$$

Substituting this relation back into Eq. (20) leads to

$$f_{Y(t)}(i\delta) = \lambda t \sum_{n=1}^{\infty} f_{N(t)}(n-1) \sum_{j=1}^{i} \frac{j}{i} f_X(j\delta) f_X^{(n-1)}((i-j)\delta)$$

Changing the order of summation in the above equation leads to

$$f_{Y(t)}(i\delta) = \lambda t \sum_{j=1}^{i} \frac{j}{i} f_X(j\delta) \sum_{n=1}^{\infty} f_X^{(n-1)}((i-j)\delta) f_{N(t)}(n-1) \tag{22}$$

From the definition of the compound sum given in Eq. (18), the second summation term is equal to $f_{Y(t)}((i-j)\delta)$. Thus, the final recursive expression is obtained as

$$\boxed{f_{Y(t)}(i\delta) = \lambda t \sum_{j=1}^{i} \tfrac{j}{i} f_X(j\delta) f_{Y(t)}((i-j)\delta)}, \quad (\text{for } i \geqslant 1) \tag{23}$$

This relation is also known as Panjer's recursion in the insurance literature, which was applied to reliability engineering problems by Noortwijk and van der Weide [30]. This expression can be easily programmed in a spreadsheet package.

The original derivation of the formula was based on the method of the probability generating function [31, Chapter 5]. In contrast, the derivation in this Section is based on probabilistic arguments, which provide more insight about the properties of the compound process. Therefore, a notable contribution of this paper is to present a clear and detailed exposition of the recursion method that is not available in the earthquake engineering literature.

5. Superposition properties of marked and compound Poisson processes

5.1. General

So far in the paper, the analysis of a single process has been discussed. However, in a realistic seismic risk analysis, a structure is likely to be affected by multiple seismic sources, each modelled as a marked Poisson process. Therefore, the overall seismic risk is a combination or probabilistic sum (or convolution) of the damage cost resulting from each seismic source. A naive way to solve this problem would be to calculate the distribution of cost using Eq. (23) for each of the m sources, and then compute another m-fold convolution to obtain the distribution of the total cost. Obviously, this is not an attractive proposition for the analysis. Furthermore, source-by-source cost analysis may not be possible, as the seismic hazard curve is typically computed by a software which provides only an aggregated curve combining the effect of all sources.

This paper shows that an elegant, single-step solution of this problem can be found using the superposition properties of the marked and compound Poisson processes, which are not well known to the seismic community, or at least not utilized by them thus far. In other words, by a single application of the recursion equation, Eq. (23), total damage cost resulting from multiple seismic sources can be computed.

Suppose an ith seismic source is modelled as a marked Poisson process, $(N_i(t), X_i)$, in which $(N_i(t); \lambda_i)$ is an HPP with rate λ_i and X_i is the mark with a distribution $F_{X_i}(x)$. The compound process associated with this source is defined as $Y_i(t) = \sum_{k=1}^{N_i(t)} X_{ik}$. Thus, the superposition principle means that these m processes can be replaced by a single equivalent HPP, $(N_{sum}(t), \lambda_{sum})$, a single marked HPP, $(N_{sum}(t), Z)$ and a single compound process, $Y_{sum}(t)$. The occurrence rate of the aggregated process is given as

$$\lambda_{sum} = \sum_{i=1}^{m} \lambda_i. \tag{24}$$

The parameters of all the three equivalent processes are summarized in the following Section.

5.2. Superposition of marked processes

In the superimposed marked HPP, the mark, Z, represents the ground motion intensity of an earthquake generated by any one of m marked processes, which has a mixed distribution given as

$$F_Z(z) = \frac{1}{\lambda_{sum}} \sum_{i=1}^{m} \lambda_i F_{X_i}(z) \tag{25}$$

The aggregated hazard curve can be defined similar to Eq. (12) as

$$H_{sum}(z) = \lambda_{sum} \overline{F}_Z(z) = \lambda_{sum} - \sum_{i=1}^{m} \lambda_i F_{X_i}(z), \quad (H_{sum}(0) = \lambda_{sum}) \tag{26}$$

Thus, the distribution of Z can be directly obtained from the hazard curve as

$$F_Z(z) = 1 - \frac{1}{\lambda_{sum}} H_{sum}(x) \tag{27}$$

Following Eq. (14), the distribution of annual maximum value is given as

$$F_{Z_{max}}(x; 1) = e^{-H_{sum}(x)} \tag{28}$$

It is very important to note that the distribution of intensity of a single event, Z in Eq. (27), is distinct from the distribution of annual maximum value, $Z_{max}(x; 1)$. given by Eq. (28).

5.3. Superposition of compound processes

The aggregated compound process is defined as $Y_{sum}(t) = \sum_{i=1}^{m} Y_i(t)$, which is equivalent to a single compound process associated with the marked Poisson process, $(N_{sum}(t), Z)$, in the following way

$$Y_{sum}(t) = \sum_{k=1}^{N_{sum}(t)} Z_k, \qquad (N_{sum}(t) > 0) \tag{29}$$

The probability distribution of $Y_{sum}(t)$ can be computed from the recursion equation, Eq. (23).

6. Application to seismic risk analysis

This Section clearly lays out the steps to apply the analytical results presented in the previous Sections for the evaluation of seismic risk.

6.1. Probabilistic seismic hazard analysis (PSHA)

PSHA is a well-established methodology for estimating the ground-motion hazard at a site of interest by considering contributions of all sources of seismicity.

The earthquake occurrence rate associated with an ith source is denoted as λ_i and the ground motion parameter as X_i. The distribution of X_i is determined as functions of seismicity parameters, e.g., the source distance, R, and the earthquake magnitude, M. Such functions are also referred to as the ground motion prediction equations. Without going into details of PSHA, the distribution of X_i is derived using the total probability theorem as

$$F_{X_i}(x) = \int_M \int_R \mathbb{P}[X_i \leqslant x | m, r] f_R(r) f_M(m) \mathrm{d}r\, \mathrm{d}m \tag{30}$$

The main outcome of PSHA is an overall seismic hazard curve, $H_{sum}(x)$, aggregating the effect of all sources, as defined by Eq. (26). In practice, computer programs are available to generate the hazard curve for any given site.

The distribution of ground motion intensity can be directly obtained by normalizing the seismic hazard curve, as given by (27). This distribution is a starting point of the damage cost analysis.

6.2. Structural response analysis

The next task in seismic risk analysis is to estimate the building response in terms of a suitable parameter, such as the maximum inter-story drift (D).

The structural response is generally given as a conditional distribution, $F_{D|Z}(d|z)$, of the drift given a ground motion intensity level. It is typically estimated by dynamic analyses of the structure subjected to representative earthquakes. The overall distribution of drift is obtained from the total probability theorem as

$$F_D(d) = \int_Z F_{D|Z}(d|z) \,\mathrm{d}F_Z(z) \tag{31}$$

6.3. Damage cost due to a single seismic event

The building damage cost, C, given the occurrence of an earthquake is a function of the extent of structural damage (D), which

Table 1
The relation of the building damage factor to the maximum inter-story drift [25].

Level	Drift Ratio (%)	Damage State	Range of Damage Factor (%)
1	< 0.2	none	0
2	0.2–0.5	slight	0–1
3	0.5–0.7	light	1–10
4	0.7–1.5	moderate	10–30
5	1.5–2.5	heavy	30–60
6	2.5–5.0	major	60–100
7	> 5.0	destroyed	100

in turn depends on the earthquake intensity. There are different ways to estimate the distribution of C.

In the first method, the damage cost is derived as a function of the drift using experience based rules, such as those given in Table 1 [25, Table 3.5, p. 48]. Note that this information is based on extensive research undertaken by the Federal Emergency Management Agency [32]. In this Table, the drift is related to the *damage factor*, which is a fraction of the total cost of replacing the building. By a combined use of the drift distribution, $F_D(d)$, and Table 1, a discrete distribution of the damage factor, $F_C(c)$, can be constructed.

The second approach is to relate the drift to the cost by an empirical relation, such as [3]:

$$\ln C = a + b \ln D + \epsilon$$

where ϵ is the error of regression. Such relations are generally developed based on simulation results or damage data from past earthquakes.

A more detailed approach is based on the concept of seismic vulnerability, which is a distribution of the cost conditioned on the ground motion parameter, i.e, $F_{C|Z}(c|z)$. The vulnerability function is typically estimated by structural dynamic simulations combined with a sophisticated model of damage of different elements of the building [19]. Once the vulnerability model is given, the final distribution of the cost per seismic event can be obtained by the total probability theorem as

$$F_C(c) = \int_Z F_{C|Z}(c|z)\, \mathrm{d}F_Z(z) \qquad (32)$$

6.4. Total cost of seismic damage

The final step is to evaluate the distribution and moments of the total damage cost, $Y_{sum}(t)$, over the service life of the structure. For this purpose, the recursion equation, Eq. (23), is solved with the distribution of C in a discrete form. The seismic risk is calculated as the expected cost of damage: $\mathbb{E}[Y_{sum}(t)] = \mu_C \lambda_{sum} t$ where μ_C is expected losses per event.

The discretization of the distribution of C is not an onerous task. Rather, it is natural to represent the cost in a convenient unit of thousand or hundred dollars. The reason is that a large uncertainty associated with damage induced by an earthquake precludes the estimation of the cost in a very fine details or small units. A continuous distribution can be discretized by any of the methods illustrated by Noortwijk and van der Weide [30].

6.4.1. Remarks

To determine the distribution of C, the starting point must be the distribution of Z, not that of Z_{max}. The reason is that Z_{max} is a maximum value of several possible earthquakes in a given time interval. Thus, using Z_{max} in the life cycle analysis would lead to the distribution of "maximum" damage cost that could incur under multiple seismic events, not the cumilative cost due to all possible events in the given interval.

Since the derivative of the hazard curve, $\mathrm{d}H(x)/\mathrm{d}x$, is related to Z_{max}, its use in the damage cost analysis is technically incorrect.

7. Practical example: seismic risk analysis

This Section presents the seismic risk analysis of a 20-storey steel moment resisting frame (SMRF) supporting an office building, which was introduced in Section 1.2 (see Fig. 1).

7.1. Seismic hazard analysis

The building is situated at the Diablo Canyon power plant (DCPP) in California. The seismic sources affecting the site are shown in Fig. 2 and resulting seismic hazard curve, $H_{sum}(x)$, is plotted in Fig. 3. The total annual frequency of occurrence of earthquakes at this site was estimated as $\lambda_{sum} = 0.106$ events per year.

The ground motion parameter, Z, is the spectral acceleration at the first period (4 s) of the structure. Its distribution, $F_Z(z)$, is calculated from the hazard curve using Eq. (27) displayed in Fig. 8.

The probability distribution of annual maximum spectral acceleration is calculated using Eq. (28) shown in Fig. 9. This distribution has an atom of 0.899 at $Z = 0$. In other words, 0.899 is the probability of no occurrence of any earthquake in any one year period.

7.2. Probabilistic response analysis

Bazzurro and Cornell [33] evaluated the nonlinear seismic response of the building and showed that the maximum interstory drift, D, can be modelled by a lognormal distribution with the logarithmic mean given as (at 2% damping):

$$\mu_{\ln D} = -2.32 + 0.7\ \ln(Z) \qquad (33)$$

The logarithmic standard deviation of the drift was estimated as $\sigma_{\ln D} = 0.37$. The probability distribution of drift, D, was computed using Eq. (31), and the resulting POE curve of the drift is plotted in Fig. 10.

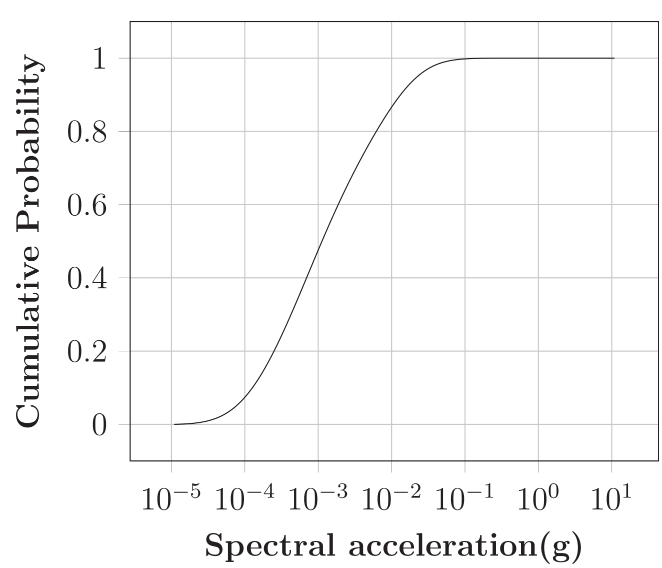

Fig. 8. Cumulative probability distribution of the spectral acceleration of any random earthquake.

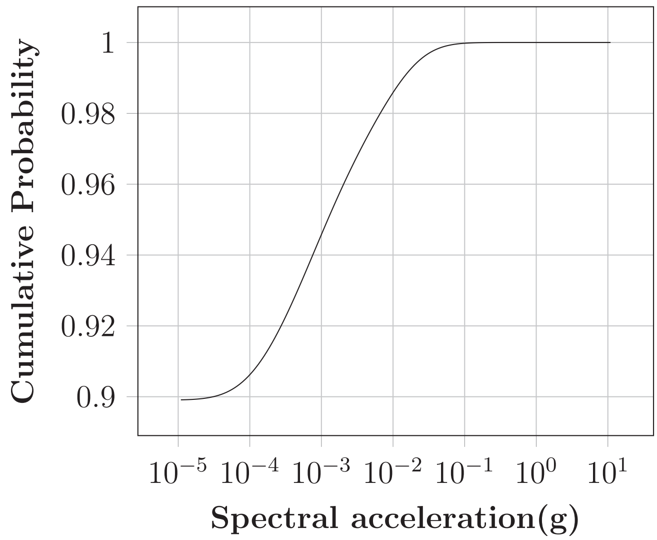

Fig. 9. Cumulative probability distribution of annual maximum spectral acceleration.

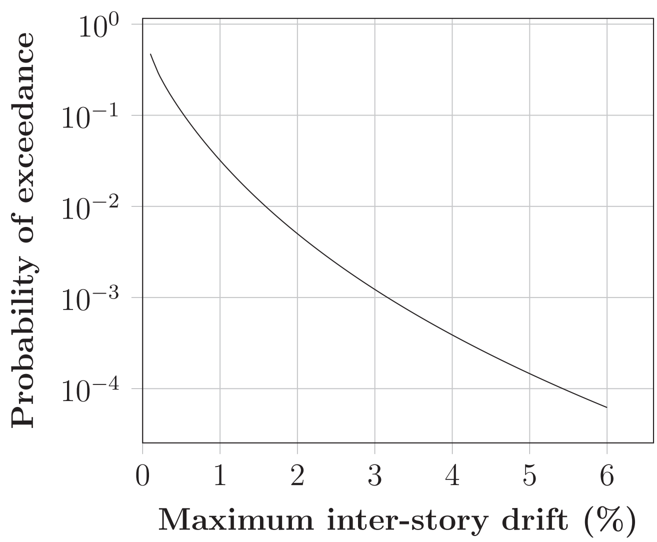

Fig. 10. Probability of exceedance (POE) of maximum inter-story drift for a 20-story SMRF structure.

7.3. Normalized damage cost due to a single earthquake

The damage cost, C, is presented in a normalized form known as the damage factor. The distribution of C is constructed using the drift POE curve given in Fig. 10 and the relationship between the drift and the damage factor given in Table 1.

Given that a range of drift, $d_1 < D \leqslant d_2$, results in the damage factor in the range, $c_1 < C \leqslant c_2$, the corresponding probabilities can be related as, $\mathbb{P}[c_1 < C \leqslant c_2] = \mathbb{P}[d_1 < D \leqslant d_2]$. For example, the drift ratio in the range of 0.7% to 1.5% can result the damage factor in the range of 10% to 30%, as shown in Table 1. From the distribution given in Fig. 10, $\mathbb{P}[0.7\% < D \leqslant 1.5\%] = 0.0529$, such that $\mathbb{P}[10\% < C \leqslant 30\%] = 0.0529$.

It was decided to discretize the damage factor between 0 and 100% in increments of 5%. The total probability in any interval, $\mathbb{P}[c_1 < C \leqslant c_2]$, was equally distributed among all the discretized values contained in the interval. For example, the damage factor in the range 10% to 30% has 6 discrete values and each of them are assigned an equal probability of 0.0529/6 = 0.0132. With this kind of judgement, the probability mass function of the damage factor, $f_C(c)$, was constructed, as shown in Fig. 11.

7.4. Total cost of seismic damage

The final step is to compute the distribution of the total cost (in terms of the damage factor) over a 50 year service life of the structure using the recursion equation, Eq. (23), with $\lambda_{sum} = 0.106$ and $t = 50$ year. The result of the computation is shown in Fig. 12.

The expected value of the damage factor, i.e., the seismic risk, is estimated as 16% over a 50 year period. There is roughly 33% probability that the damage factor would exceed the mean value. The 95th percentile of the damage factor is 55%, which is an upper bound estimate of the the damage cost.

It is clear that the average value of damage cost is not a realistic estimate of actual potential of losses, as there is a high probability the actual damage cost would exceed the average value. With the full distribution of cost a more realistic upper bound can be estimated, which is the main advantage of the proposed approach.

8. Summary and conclusions

Although there is a substantial body of literature on the seismic life cycle analysis, the evaluation of probability distribution of total cost of seismic damage has not been tackled in a rigorous manner. The paper attempts to fill in this gap by presenting a systematic development of stochastic modeling of seismic risk analysis that

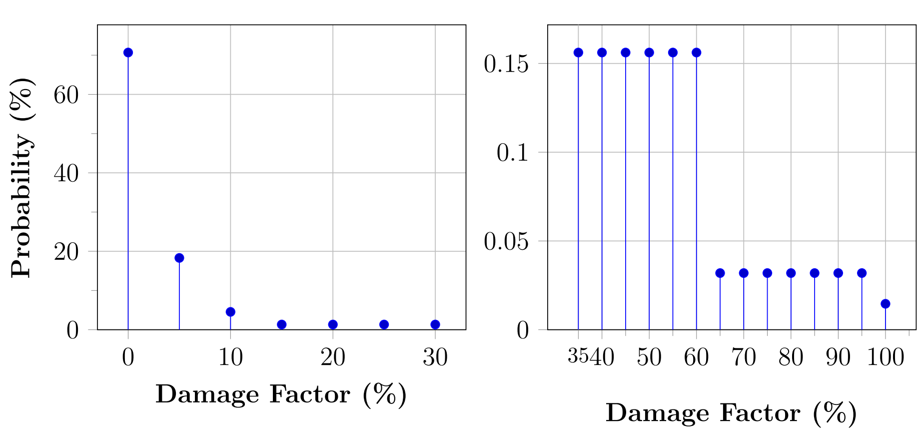

Fig. 11. Distribution of the normalized cost (or damage factor) due to a single earthquake.

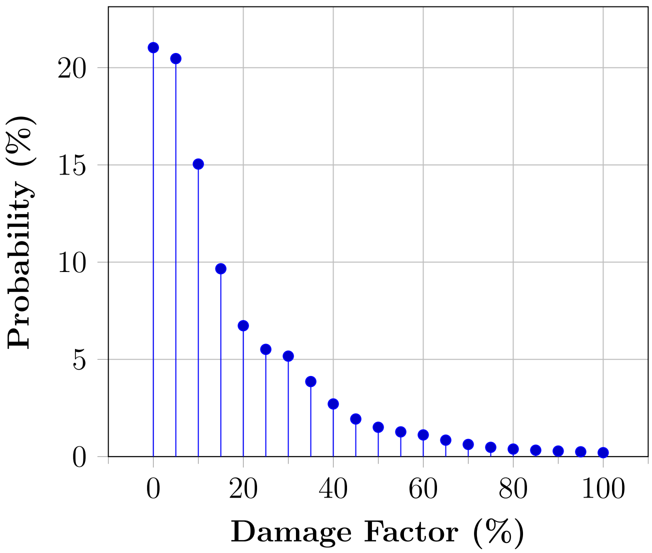

Fig. 12. Normalized seismic damage cost over a 50 year period.

leads to the distribution of cost without a recourse to a simulation-based method.

Since the pioneering work of Cornell [5], the seismic reliability analysis is based on the distribution of the maximum of a marked Poisson process. The present paper extends this thinking further to show that the total cost of seismic damage is a compound Poisson process, for which an analytical solution is possible. The paper shows how to use the superposition property of the compound Poisson process to account for the effect of multiple seismic sources in a single step of the analysis.

The probability distribution of total damage cost is computed by a discrete recursion equation. The proposed approach does not require any new information, other than what is contained in a standard seismic analysis, namely, hazard curve, structural response function, and the vulnerability or cost function.

A practical example is presented to illustrate all the steps of the proposed method. The expected damage factor of a 20-story frame structure is estimated as 16% of the replacement cost and the 95th percentile of damage factor as 55%.

The study brings the following new elements to the seismic damage cost analysis:

- A rigorous formulation of the life cycle cost of damage in terms of the compound renewal/Poisson process.
- A clear and self-contained exposition of the recursion method based on probabilistic arguments, which is is amenable to structural reliability community.
- The simplification of the damage cost analysis involving multiple seismic sources via the principle of superposition of compound Poisson processes.
  It must be emphasized that the recursion formula by itself has little practical value. It is the combination of the superposition property with recursion formulas, as proposed in the paper, leads to an efficient and practical solution of the problem.
- An illustration of a conceptual inaccuracy due to an inadvertent use of the hazard curve or the distribution of annual maximum ground motion intensity in the ldamage cost analysis.

Acknowledgement

The authors gratefully acknowledge the financial support for this study provided by the Natural Science and Engineering Council of Canada (NSERC) and the University Network of Excellence in Nuclear Engineering (UNENE).

References

[1] Whitman RV, Biggs JM, Brennan III JF, Cornell CA, de Neufville RI, Vanmarcke EH. Seismic design decision analysis. J Struct Div, ASCE 1975;101 (ST5):1067–84.
[2] Whitman RV, Cornell CA. Design. In: Lomnitz C, Rosenblueth E, editors. Seismic risk and engineering decisions. Amsterdam: Elsevier Scientific Publishing Company; 1976. p. 339–80 [Ch. 10].
[3] Cornell CA, Jalayer F, Hamburger RO, Foutch DA. Probabilistic basis for 2000 sac federal emergency management agency steel moment frame guidelines. ASCE J Struct Eng 2002;128(4):526–33.
[4] Whitman R. Basic concepts and important problems. In: Hansen R, editor. Seismic design for nuclear power plants. Cambridge, MA: The M.I.T. Press; 1970. p. 1–68.
[5] Cornell CA. Engineering seismic risk analysis. Bull Seismol Soc Am 1968;58 (5):1583–606.
[6] Pandey MD, Wang Z, Cheng TJ. Stochastic renewal process models for life cycle cost analysis. ICASP; 2015.
[7] Gupta A. Seismic demands for the performance evaluation of steel moment resisting frame structures [Ph.D. thesis]. Stanford University; 1999.
[8] Wang Z. Seismic risk analysis of nuclear energy facilities [Ph.D. thesis]. Waterloo, ON, Canada: University of Waterloo; 2015.
[9] Abrahamson NA, Silva WJ. Empirical response spectral attenuation relations. Seismol Res Lett 1997;68(1):94–127.
[10] Esteva L. Seismic risk and seismic design decisions. In: Hansen R, editor. Seismic design for nuclear power plants. Cambridge, MA: The M.I.T. Press; 1970. p. 142–82.
[11] Rosenblueth E. Optimum design for infrequent disturbances. ASCE J Struct Div 1976;102(ST9):1807–23.
[12] Rosenblueth E. Optimum design to resist earthquakes. ASCE J Eng Mech Div 1979;105(1):159–76.
[13] Liu M, Wen YK, Burns SA. Life cycle cost oriented seismic design optimization of steel moment frame structures with risk-taking preference. Eng Struct 2004;26:1407–21.
[14] Krawinkler H, Zarcian F, Median RA, Ibarra LF. Decision support for performance-based design. Earthquake Eng Struct Dyn 2006;35:115–33.
[15] Rackwitz R. Optimization and risk acceptability based on the life quality index. Struct Saf 2002;24:297–331.
[16] Rackwitz R, Joanni A. Risk acceptance and maintenance optimization of aging civil engineering infrastructures. Struct Saf 2009;31(3):251–9.
[17] Streicher H, Joanni A, Rackwitz R. Cost-benefit optimization and risk acceptability for existing, aging but maintained structures. Struct Saf 2008;30(5):375–93.
[18] Goda K, Hong HP. Optimal seismic design considering risk attitude, societal tolerable risk level, and life quality criterion. ASCE J Struct Eng 2006;132 (12):2027–35.
[19] Porter KA, Beck JL, Shaikhutdinov R, Au S, Mizukoshi K, Miyamura M, et al. Effect of seismic risk on lifetime property value. Earthquake Spectra 2004;20 (4):1211–37.
[20] Cha EJ, Ellingwood BR. Seismic risk mitigation of building structures: the role of risk aversion. Struct Saf 2013;40:11–9.
[21] Williams R, Gardoni P, Bracci J. Decision analysis for seismic retrofit of structures. Struct Saf 2009;31:188–96.
[22] Koduru SD, Haukaas T. probabilistic seismic loss assessment of a vancouver high-rise building. J Struct Eng 2010;136(3):235–45.
[23] Martins L, Silva V, Marques M, Crowley H, Delgado R. Development and assessment of damage-to-loss models for moment-frame reinforced concrete buildings. Earthquake Eng Struct Dyn 2016;45:797–817.
[24] Mahmaoud H, Cheng G. Framework for lifecycle cost assessment of steel buildings under seismic and wind hazards. J Struct Eng ASCE 2017;143 (3):1–15.
[25] Kang YJ, Wen YK. Minimum life-cycle cost structural design against natural hazards Tech. Rep. SRS 629. University of Illinois at Urbana-Champaign; 2000.
[26] Gosh J, Padgett JE. Probabilistic seismic loss assessment of aging bridges using a component-level cost estimation approach. Earthquake Eng Struct Dyn 2011;43:1743–61.
[27] Deco A, Bochhini P, Frangopol DM. A probabilistic approach for the prediction of seismic resilience of bridges. Earthquake Eng Struct Dyn 2013;42:1469–87.
[28] Tijms HC. A first course in stochastic models. New York: John Wiley & Sons; 2003.
[29] Taylor HM, Karlin S. An introduction to stochastic modeling. 3rd ed. San Diego, CA: Academic Press; 1998.
[30] Noortwijk JM, van der Weide JAM. Applications to continuous-time processes of computational techniques for discrete-time renewal processes. Reliab Eng Syst Saf 2008;93:1853–60.
[31] Panjer HH. Operational risks: modeling analytics. Hoboken, New Jersey: Wiley-Interscience; 2006.
[32] FEMA. A benefit-cost model for the seismic rehabilitation of buildings Tech. Rep. FEMA 227. Washington DC, USA: Federal Emergency Management Agency; 1992.
[33] Bazzurro P, Cornell CA. Vector-valued probabilistic seismic hazard analysis. In: Proc 7th U.S. national conf on earthquake engineering, Boston, MA; 2002. p. 1–8.