16th World Conference on Earthquake, 16WCEE 2017
Santiago Chile, January 9th to 13th 2017
Paper N° 3244
Registration Code: S-Y1463089227

THE P IN PSHA

X. Tao[1], Z. Tao[2]

[1] *Professor, Institute of Engineering Mechanics, CEA, and School of Civil Engineering, Harbin Institute of Technology, China, taoxiaxin@aliyun.com*

[2] *Assoc. Professor, Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, CEA, China, taozr@foxmail.com*

Abstract

Probabilistic seismic hazard assessment (PSHA) has been developed and widely applied for about half a century. Meanwhile, it has been queried and debated unceasingly, especially after some highly destructive earthquakes occurred in areas with relative low hazard on PSHA maps. It was asked "where does probability come in to play?" even "is PSHA science?" Since the data for statistics in PSHA is not enough for most parts of the world, some expert experience and judgment are practically involved in almost every PSHA project. The key point to understand the P in PSHA is emphasized in this paper as it is not only impossible but also unnecessary to take into account the correlation between earthquakes in the given future time period. Poisson distribution is derived just based on the independence assumption and the fact that the occurrence of destructive earthquake is rare. A reasonable consideration is to take a statistic unit as large in space and as long in time as possible. For the former, the limitation is if a larger area is considered to develop G-R relation, the hazard may be underestimated at the site of interest where seismicity is higher than other locations in the area. For the latter, the limitation comes from the fluctuation of seismicity and the completeness of earthquake catalog. Chinese scientists introduced an approach of believable magnitude and two-rank delineation of potential source area to deal with these problems. The most significant step in the two-rank procedure is to assign the seismicity parameters estimated for the first rank unit into the sub-source areas in the second rank by weighting factors. It is quite difficult to evaluate the weighting factors by means of a comprehensive understanding of the earthquake occurrence from seismic, tectonic and crust dynamic evidences. Some artificial intelligence tools, such as Patter Recognition, Artificial Neural Network are suggested in the paper to avoid subjective judgment in the evaluation of the factors. The conclusion is that we are all on the way to a confident SHA, we have some data and know something on the occurrence of strong earthquakes, but cannot really predict it in a strict scientific way even for long-term, that fact is the reason of a P in PSHA.

Keywords: probability; PSHA; independent; Poisson; uncertainty

1. Introduction

Seismic zoning map plays a role of governmental policy for earthquake fortification in most of earthquake-prone countries, and as a national standard in some countries like China. Probabilistic seismic hazard assessment (PSHA) has been developed and widely adopted in compiling of hazard map and in evaluating of major project sites for about half a century. Meanwhile, it has been queried and debated unceasingly, especially after some highly destructive earthquakes occurred in areas with relative low hazard on PSHA maps. In the criticisms, it was asked “where does probability come in to play?” even “is PSHA science?” [1, 2, 3]. The formulas were also challenged as inaccurate and lead to systematic errors [4, 5]. As researchers working on PSHA for decades, the authors would like to explain the fundamentals of seismic hazard assessment (SHA), what the advantage of PSHA is, the assumption in Poisson model, issues in estimating seismicity parameters for the P from earthquake data by their personal understandings, and to mention some results of their group on the uncertainty correction of the P and test of hazard map of PSHA, in this paper.

2. Fundamentals of SHA and PSHA

Seismic hazard mapping in the world initiated at the end of the 19th century in Russia, mainly from observed Intensity data. So the map was actually maximum observed Intensity map. During the mid of the last century, long-term earthquake prediction was involved in the zoning, it was from a comprehensive understanding of earthquake occurrence mostly with seismological and tectonic consideration. Afterwards, some researchers mainly worked on statistics of historical data e. g. [6, 7], while some others emphasized on the tectonic cause of earthquakes e. g. [8]. The fundamentals of SHA to estimate strong earthquake occurrence from regional seismicity and tectonic condition was summarized as the following two [9]: (1) strong earthquake may occur at a place where a destructive shock occurred with the similar magnitude, (2) strong earthquake may occur at a place where the tectonic condition is similar with another place where a destructive shock occurred. The above mentioned “may” shows uncertainty of the “occur” from the knowledge learned from the past earthquakes themselves, thus the most difficult work in SHA is to manage this uncertainty, to define the “place” in which space-time area for the first fundamental, and to characterize the tectonic conditions and their “similarity” for the second one. Up to now, the common view on locations and activities of potential sources of tectonic earthquakes is that they may be many and different in kind, and may not even be well known. In some regions, it is not possible to correlate past activity with known tectonic features.

Probabilistic approach is powerful to manage uncertainty, but is incapable to reduce it. Therefore, Cornell did not attempt, when he built the foundation and the frame of PSHA, to turn away from these two fundamentals, he tried to adopt all the pertinent data and professional judgments of those trained in seismology and geology [10]. For purpose of the development, he emphasized to express seismic risk in terms of return periods from the requirement for seismic design of engineering project, so that engineers can make a trade-off costly between higher resistances and higher risk of economic losses. Engineers must consider the performance of the project under moderate as well as strong motions, and how quickly the risk decreases as the intensity increases, so intensity versus average return period is far more useful than single numbers as the "expected lifetime maximum" or "50-year" intensity. The advantage doesn’t mean that PSHA can always predict hazard for a site or region more truly than the deterministic seismic hazard assessment (DSHA). Actually, results of PSHA were even modified by results of DSHA in some cases, e. g. in the Yucca Mountain project [5]. Therefore, PSHA in some countries is considered on a basis of earthquake long term prediction, e. g. in Japan [14, 15]. In order to simulate near fault ground motion field appropriately, an idea is proposed to combine the PSHA and DSHA approaches for the next generation map, in which scenario earthquake is suggested as the link between the two [11].

3. What is the hypothesis in Poisson model

The authors believe that most problems about whether to assume that the probability of earthquake is constant or varies with time come from Poisson assumption. In fact, Poisson distribution is derived just based on the independence assumption and the fact that the occurrence of destructive earthquake is rare, although some researchers emphasize some additional assumptions such as the exponentially distributed of inter-arrival times. Let us start from the simplest case, the P (exceeding probability) of ground motion amplitude Y lager than a given value y at a site is just caused from two earthquakes, E_{ij} with magnitude M_{j} in the i_{th} potential source area and E_{kl} with the same meaning but in the k_{th} source. It can be computed as

$$P(Y>y)=P(Y>y|E_{ij})\cdot P(E_{ij})+P(Y>y|E_{kl})\cdot P(E_{kl})-P(Y>y|E_{ij}\cap E_{kl})\cdot P(E_{ij}\cap E_{kl})$$

(1)

The three conditional probabilities at the right side of the above equation depend on the regional ground motion attenuation relationship, and the spatial correlations between the site and the source areas. The occurrence probabilities of E_{ij} and E_{kl} are actually quite difficult to estimate in PSHA, and are going to be discussed later. Here the fact that the probability of simultaneity of E_{ij} and E_{kl} is much more difficult to estimate is emphasized, since estimation of the conditional probability of E_{ij} occurrence given E_{kl} occurred is much more difficult than estimation of occurrence probability of E_{kl}. If independency between the two earthquakes is accepted, Eq. (1) can be simplified as

$$\begin{aligned}P(Y>y)&=P(Y>y|E_{ij})\cdot P(E_{ij})+P(Y>y|E_{kl})\cdot P(E_{kl})-P(Y>y|E_{ij})\cdot P(E_{ij})\cdot P(Y>y|E_{kl})\cdot P(E_{kl})\\&=1.0-\left[\left(1.0-P(Y>y|E_{ij})\cdot P(E_{ij})\right)\cdot\left(1.0-P(Y>y|E_{kl})\cdot P(E_{kl})\right)\right]\end{aligned}$$

(2)

It makes PSHA much easier to assume independency between each pair of earthquakes, since there must be some quakes, near or far, strong or weak, which should be taken into account. In actual case, Eq. (2) can expressed for those quakes as

$$P(Y>y)=1.0-\prod_i\prod_j\left(1.0-P(Y>y|E_{ij})\cdot P(E_{ij})\right)$$

(3)

in which, the first P term in brackets at the right side means the probability of Y greater than y given E_{ij} occurred, $P(E_{ij})$ the probability of E_{ij} occurrence, i and j must take all possible events in a given time period t into account.

Firstly, we divide the period t into N time intervals end to end, each interval with a length $dt=t/N$. In each interval, there are two possibilities for earthquake occurring or no occurring. The probability of occurrence, P in every interval can be considered as the same, since their lengths are same. Under the independent hypothesis, the events in N intervals form an independent repeated trial. From the Bernoulli formula, the probability of earthquake occurring in n intervals and no quake in N-n intervals among the N trials is $P^n(1\text{-}P)^{N\text{-}n}$. We just care how many occurrence intervals there can be, rather than which one they are, so there are totally C_N^n combinations, the probability should be

$$\begin{aligned}P(n)&=C_N^nP^n(1-P)^{N-n}=\frac{N(N-1)\cdots(N-n+1)}{n!}P^n(1-P)^{N-n}\\&=\frac{NP(N-1)P\cdots(N-n+1)P}{n!}\cdot\frac{(1-P)^{-\frac{1}{P}(-NP)}}{(1-P)^n}\end{aligned}$$

(4)

When N is very large and P is very small, it is true in nature since earthquake occurrence is seldom, NP in t reaches the mean μt (small positive), and all of $(N\text{-}1)P$ to $(N\text{-}n\text{+}1)P$ reach μt as well, meanwhile $(1-P)^n \to 1.0$ and $(1-P)^{-\frac{1}{P}} \to e$. Therefore, the probability of n earthquakes occurring is as following Poisson formula

$$P(n) \approx \frac{(\mu t)^n}{n!} e^{-\mu t} \tag{5}$$

Substitute Eq. (5) into Eq. (3), and summary all probabilities for 0 to infinity of occurrence by total probability formula, we can get

$$\begin{aligned}
P(Y > y) &= 1.0 - \prod_i \prod_j \sum_{n=0}^{\infty} \left(1.0 - P\left(Y > y \middle| E_{ij}\right)^n \cdot P_{ij}(n) \right) \\
&= 1.0 - \prod_i \prod_j \sum_{n=0}^{\infty} \left(1.0 - P\left(Y > y \middle| E_{ij}\right)^n \cdot \frac{(\mu_{ij} t)^n}{n!} e^{-\mu_{ij} t} \right) \\
&= 1.0 - \prod_i \prod_j e^{-\mu_{ij} t} \cdot \sum_{n=0}^{\infty} \left(1.0 - P\left(Y > y \middle| E_{ij}\right)^n \cdot \frac{(\mu_{ij} t)^n}{n!} \right) \\
&= 1.0 - \prod_i \prod_j e^{-\mu_{ij} t} \cdot e^{\mu_{ij} t \left(1.0 - P\left(Y > y \middle| E_{ij}\right)\right)} \\
&= 1.0 - \prod_i \prod_j e^{-\mu_{ij} t \cdot P\left(Y > y \middle| E_{ij}\right)} \\
&= 1.0 - e^{-\sum_i \sum_j \mu_{ij} \cdot t \cdot P\left(Y > y \middle| E_{ij}\right)}
\end{aligned} \tag{6}$$

where, $P(Y>y|E_{ij})$ depends on attenuation models [12, 13], will not be discussed further.

It is clear that the hypothesis of Poisson model is just the independency between earthquakes in the given future time period, and the seldom nature of earthquake occurrence. The unique parameter μ is the annual mean occurrence rate in that period, not about any time period before or after that. This conclusion could be an answer to the question if the rate is constant or variable in Poisson model. Following it, the authors would like to point out the discussion by Cornell, “If the engineer and the seismologist are prepared to make an assumption about the time dependence of the average occurrence rate, other than that of constant in time, a minor modification in the method suffices to account for this non-homogeneity in time” [10], may not be on the above formulas, but on the estimation of the mean rate. It means one does not have to modify the method, but just take another value. Different rate value can reflect the estimation that a zone of recent past activity is less likely to be the source of the next strong earthquake than a previously active zone which has been relatively quiet for some time. Furthermore, the more general models such as renewal process or Markov process can be adopted also only on the rate estimation, not be combined into the above formulas. The Eq. (59) in the funding paper of PSHA [10] is unnecessary, since the integration result must be μt if the field of the integration is 0 to t. Cornell mentioned numerous reasons to accept Poisson model even when more accurate theoretical models become available [10], the authors like to add one more that for engineering purpose, it is not only impossible but also unnecessary to take into account the correlation between earthquakes in studying losses due to a succession of moderate quakes or cumulative damage due to two or more major shocks.

4. How to estimate the seismicity parameters — a suggestion by Chinese scientists

In nature, parameters concerning with the *P* must come from statistical data, and the most principal contradiction in PSHA comes from the detail hazard mapped for construction and development planning by means of statistics in quite large region. PSHA is deemed as erroneous, since the data for statistics in it is not enough for most parts of the world, especialy data of large earthquake, some expert experience and judgment are practically involved in almost every PSHA project. Hazard maps depends on their makers' considerations. This dependence can be found by comparing maps of the same area made by different groups, which can predict hazards differing by factors of three to four in some cases. These differences show some of the uncertainties that make assessing the performance of hazard maps crucial. The authors believe that the key point in PSHA is to carry out the statistical analysis for seismicity parameters in what "space-time area". To predict large earthquake with magnitude 8 or more, one should watch a quite large area, not from statistics in a small area. Even in subduction zone, like east sea of Japan with so many earthquake data, hazard of a shock 9.0 is still not be recognized if statistics is carried just on the data in small areas such as off Miyagi and/or off southern Sanriku. In "off Miyagi" region, six events repeated in the past 200 years with the average interval of 37.1 years, which resulted the occurrence probability of 99% in 30 years after 2009, the highest value of the all areas predicted off the Pacific coast of Japan, but with magnitude just 7.5 for independent occurrence, or 8.2 for correlated occurrence with its neighbor area [14, 15]. However, an earthquake with Mw8.8 in 1988-2018 may be estimated from analysis in the whole subduction-zone of East Japan [16]. Clearly, the larger spatial range is, the more earthquake data will be, but the difference between the future seismicities at various locations in it will not be shown.

In order to cope with this problem, Chinese scientists suggested a two-rank scheme of potential source areas [17]. Seismic province or zone is taken as source rank A to indicate a large region, with constitutes a tectonic unit and encompass adequate earthquake data for statistical analysis, so that seismicity parameters, such as upper bound magnitude, annual recurrence rate, *b* value etc., could be evaluated in general. Rank A source could be defined with considerations on seismicity data (destructive and small events), tectonic data (active fault, basin, blocks), geodetic data (such as GPS), geophysical data (for the crust structure and deep structure) and geodynamic understanding. Source areas of rank B are sub-areas in the source A, to indicate the areas with its own upper bound magnitude, and their annual recurrence rates in each magnitude interval will be assigned by a set of weighting factors from those of rank A source. From Gutenberg-Richter type relation, $\mu(m)$, the annual recurrence rate of earthquake with magnitude greater than *m* is

$$\mu(m) = 10^{a-bm} \tag{7}$$

where, *a* and *b* are coefficients of the relation from a rank A source. Then μ_j, the annual recurrence rate of earthquake in the j_{th} magnitude interval can be calculated as follows

$$\mu_j = \mu(M_j - \tfrac{1}{2}\Delta M) - \mu(M_j + \tfrac{1}{2}\Delta M) \tag{8}$$

where, $M_j - \frac{1}{2}\Delta M$ and $M_j + \frac{1}{2}\Delta M$ are the lower and upper limits of the interval respectively. The annual recurrence rate in the j_{th} magnitude interval and in i_{th} rank B source area, μ_{ij} can then be obtained by

$$\mu_{ij} = \mu_j \cdot W_{ij} \tag{9}$$

where, W_{ij} is the corresponding weighting factor for the i_{th} area and the j_{th} magnitude interval, and

$$\sum_i W_{ij} = 1.0 \tag{10}$$

in which the summation just cover all rank B sources with upper bound magnitudes greater than $M_j - \frac{1}{2}\Delta M$.

It is also a reasonable consideration to take a statistical analysis in time period as long as possible. In practice, the limitation comes from the completeness of earthquake catalog and the fluctuation of seismicity history. The earthquake history available from instrumental records is usually too short, compared with the long and variable recurrence time of large earthquakes. China has the longest earthquake record in the world, but there must be many quakes not recorded in the historical time periods without monitoring networks. Chinese seismologists suggested a term as “believable period” for some magnitudes to show the situation that the earlier the bigger earthquakes may not be recorded. For example, the periods from 1561, 1604 and 1885 to now are suggested to be believable respectively for magnitude greater than 7.0, 6.0, and 5.0 in the Gansu region [18]. If one wants to fit a recurrence relation as Eq. (7), the mean occurrence rates of events with the magnitudes should be counted from data in the three periods. The seismicity fluctuation showed in regional earthquake history may also be taken into account with parameters estimated from those data in corresponding time periods, by a confident inferring of seismicity tendency in the future. Clearly, the shorter time period is taken, the fewer earthquake data will be, and the seismicity parameters will be closer to some kind of long term prediction.

5. The more complicated issue

The most significant step in the two rank procedure is to assign the seismicity parameters estimated for the rank A source into the rank B sub-source areas within by weighting factors. Then the problem turns on the weighting, it is quite difficult to evaluate the factor values from a comprehensive understand of the occurrence from seismic, tectonic and crust dynamic evidences, since it is not matured quantitative to characterize the tectonic condition and the “similarity”, and a logical approach to add up contributions of all these information is still requiring further study. As shown in Eq. (10), the upper bound magnitudes of all rank B sources are one controlling factor too. The shortness of the earthquake records can also cause hazard assessment to be biased by recent regional largest events, which in general produce high-hazard areas on maps [19]. These areas can be misleading, especially on continents where the spatiotemporal patterns of seismicity are more irregular than those at plate boundaries. The large scale numerical studies based approaches directly on historical data also have difficulty giving proper weight to the known correlation between geological structure and most seismic activity [10]. In fact, it in some degree involves the causality between the two as mentioned as the second fundamental mentioned above. Chinese geologists generalize some tectonic criterions for earthquakes with magnitude 6.0, 7.0 and 8.0 for various regions from many case studies. They may be helpful to evaluate weighting factors, but do not always work well. One example is the case of Wenchuan earthquake, a shock M8.0 occurred at Longmenshan fault where the slip rate is quite slower than that on nearby Xianshuihe fault where strong earthquake occurred quite often, therefore Longmenshan fault was assessed as low hazard. The causation is mentioned by experts after the disaster as the fault locked the east wards moving of the blocks at the eastern edge of the Qinghai-Tibet plateau where is the most active region in mainland of China. This lesson demonstrates that the criterions should be modified and developed further, and the factors in the criterions must be complicated not only themselves, but also the relations between each pair of them. For example, the slip rate and the observed maximum magnitude of event on a fault could be positive or negative factor which depends on the movements of the faulting system around it. So the earth dynamical analysis may be necessary for the weighting factors. At present, the estimation of weighting factors is naturally a inferring with incomplete data base and empirical knowledge, and could be considered as a decision-making problem with some artificial arbitrariness and personal bias. It likes a work to locate the position(s) of potential earthquake with given magnitude in the rank A source, given the number of events. From the daily knowledge, one may seek out the positions as more as possible to be conservative; however it may lead to an underestimation of hazard, since the annual recurrence rate of strong earthquake in general is quite low, and must be much lower if it is distributed into more rank B sources. Some artificial intelligence tools, such as Patter Recognition, Artificial Neural Network, are suggested to avoid subjective judgment in the evaluation of the factors.

Weighting factor must be generalized from a set of specialized factors. The approach to sum up all the factor values needs further study. Some factor values may be added together, but some should be multiplied, such as the “immunity” after a large shock. Immunity is a term for the fact that there will be no strong quake occurring in a time-spatial range after a large earthquake since energy there in the crust should be accumulated in a long time for the next shock.

6. Uncertainty correction

From the above discussion, one can see there must be a lot of uncertainty in a PSHA result, due to the lack of enough detail seismic and geological data and poor understanding of effect of focal mechanism, seismic wave transmission and the local site conditions, in which the uncertainty of attenuation relationship is dominant. The uncertainty process is one of the most focuses in the debates [20, 21]. Esteva revealed the fact that the distribution of the random error of motion attenuation was Normal distribution, and developed a correction procedure [22]. Ang and Der Kiureghian suggested a procedure to combine the uncertainty of the attenuation into inherent uncertainty of ground motion by means of total probability formula, after seismic hazard estimated from a mean attenuation relationship [12]. It is called as 'correction' now, and its rationality is queried since there are often some large increases on the corrected hazard especially at very low exceeding probability range.

Attenuation relation predicts value of a ground motion parameter for given magnitude and distance with simplifications on source mechanism, wave propagation and local site condition. Survey and analysis in several decades show the fact that there is always a quite large variation in ground motion, even if from similar magnitude, depth, distance and site condition, for whatever which motion parameter is. In general, attenuation relationship of ground motion could be represented as following formula

$$\ln Y = \ln A(M, R) + \varepsilon \tag{11}$$

where, Y is for motion parameter, M for magnitude, R for distance, $A(\bullet)$ for the attenuation function form, ε is for the difference between $\ln Y$ and $\ln A(M, R)$, called as random error with a Zero-mean Normal distribution. The only one numerical character of the distribution, standard deviation σ represents the inherent randomness, error from simplification, and validity of the f function form. It is not hopeful to reduce the σ obviously by means of adding any more variable in the attenuation relationship, or expending data base from accumulated observing data [23].

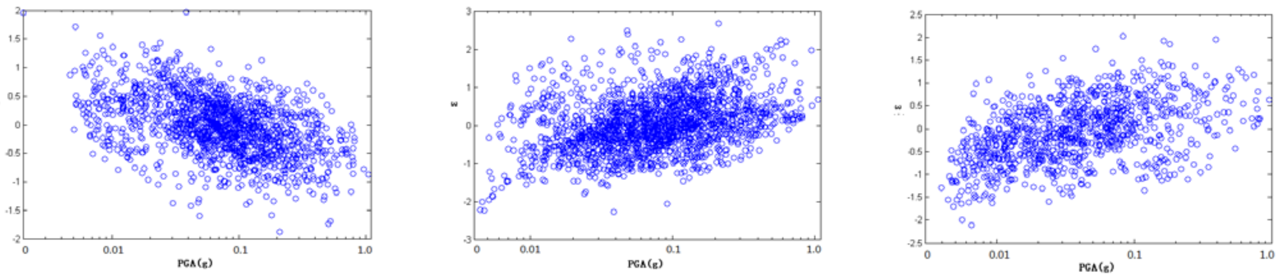
From the total probability theorem, the correction formula can be expressed as

$$P(Y_c > y) = \int_{-3\sigma}^{3\sigma} P\left(Y > ye^{-\varepsilon}\right) f(\varepsilon) d\varepsilon \tag{12}$$

where, Y_C is for the corrected ground motion parameter, Y is for the uncorrected one, $f(\varepsilon)$ is the probability density function of ε. It is obvious that the field of the integration is trunced from negtive infinity to positive infinity, since ground motion is believed be less than a upper bound.

Some researchers propose that calculation of Eq. (12) requires the knowledge of the joint probability distribution of all random variables involved. Klügel suggests that it leads to a systematic double counting of uncertainties if those in source parameters are modeled additionally and independently in the source model [4]. The authors disagree with this understanding of double counting, since the distribution $f(\varepsilon)$ is from the observed motion with well determined magnitude and measured distance, nothing about the uncertainty on location and magnitude of future earthquake. Uncertainties of the conditional probability $P(Y>y|E_{ij})$ and occurring probability $P(E_{ij})$ in Eq. (3) can be corrected independently. The latter processed by logical tree will not be discussed hereon; problems on the former to be dealt with in deep could be as follows, if the truncated range could be narrowed down further? If the distribution depends on magnitude, distance, or ground motion amplitude such as PGA? And, how to improve the uncertainty correction procedure with any new findings?

NGA project released a good data base of ground motion, and worked out some attenuation relationships, e. g. C-B (Campbell-Bozorgnia, [24]), C-Y (Chiou-Youngs, [25]) and Id (Idriss, [26]). The ε values from the three relations are calculated from the data base, and the result shows that the numerical characters do not vary with magnitude and distance obviously, but the mean values change with acceleration very clearly while σ is quite stable [27], as shown in Fig. 1.



C-B model C-Y model Id model

Fig. 1 – The distribution of ε with acceleration

An improvement to distribution of random error is suggested from the above findings as Normal distribution with same σ and subsection mean, $f_y(\varepsilon)$. From situation that a continuous function of the mean on acceleration cannot be developed from the limited data at present, a subsection correction is suggested as

$$\mathrm{P}(Y_C > y) = \int_{\mu_i - 3\sigma}^{\mu_i + 3\sigma} P(Y > ye^{-\varepsilon}) f_{\mathrm{y}}(\varepsilon) d\varepsilon \tag{13}$$

where, μ_{i} is the mean of ε in the i^{th} subsection of y. The hazard curves for a site by attenuation relation C-B as an example are shown in Fig. 2, in which the thick solid line is for uncorrected curve, the dash line for subsection corrected curve, the thin solid line for routine corrected. One can see from the figure that the expected acceleration for the low exceeding probability by subsection correction is much smaller than that by correction with mean 0, according to the decreasing of ε with acceleration increases.

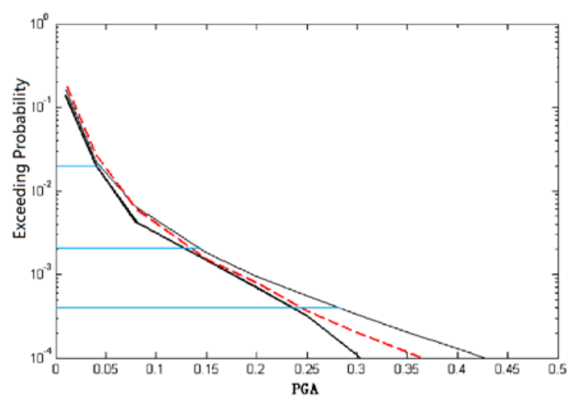


Fig. 2 – The hazard curves for an engineering site from [27]

7. Test of hazard map of China

Among all debates and queries on PSHA, the demand to test hazard map cannot be ignored, since hypothesis testing is the heart of the scientific method. Notwithstanding the difficulties, it is essential that a continuing process of serious and objective testing is conducted for the methods used to produce seismic hazard maps [2]. This would involve developing objective criteria for testing such maps by comparison to the fact that earthquakes actually occurred after the maps publication. The key point is how to manage the exceeding probability and intensity difference. Ideally, one can examine the difference between happened intensity and that on the map repeatedly for many time periods, hundreds of years, and then get a ratio to see if it is comparable. However, hazard maps in most earthquake-prone countries are generally renewed in every 10 years or so, people cannot wait so long after the map taken out of service. In fact, PSHA map is compiled by means of systematically applying to a grid of ground points, hazard at each point is assessed from the same set of regional potential source areas and seismicity parameters, same attenuation relations for a quite large territory, but independently, i. e. motion exceeding a given value at a point does not depend on the exceeding at any other point nearby. It would insure that consistent assumptions were being used for all portions of the region and among different regions [10]. Therefore, one way to test PSHA map could be to examine the intensity difference at every point on the map for a time period, the mean exceedance could be considered as the mean ratio for many periods if the points are many enough, because repeat trial in time domain is equivalent to trial in space domain. It is important to test maps by means of records in time as long as possible; the major challenge for such testing is the availability of only a relatively short earthquake shaking record. From the independent hypothesis, the hazard difference between any two sub periods in the map service time period are not taken into account, neither the relation between an earthquake and the following next one, so the exceeding probability in a given period can be converted as that in any piece of time within. For example, the service time of PSHA maps of China is 50 years; the exceeding probability in t years, P_t can be converted from the probability P_{50} by

$$P_t = 1.0 - (1.0 - P_{50})^{\frac{t}{50}} \tag{14}$$

The t value are taken as 22 and 11 years, since the two maps of China were issued in 1990 and 2001 respectively, and corresponding P_t are 4.5% and 2.3% from 10% of P_{50}. Isoseismals of 238 and 119 shallow earthquakes with M≥5 occurring in the mainland of China from 1990 to 2011, and 2001 to 2011, are respectively overlaid on the maps by means of the spatial manipulation and analysis function of Geographic Information System (GIS). The mapping parameter Basic Design Acceleration (BDA) is converted to Macro Intensity by VI, VII,VIII and IX for 0.05g, 0.1g and 0.15g, 0.2g and 0.3g, and 0.4g, so that to be comparable with the isoseismals. For some earthquakes without surveyed isoseismals, the isoseismals are generated by a set of attenuation relationships. The total areas with observed intensity at least one degree larger than the corresponding intensity on the maps are summed as 2.67% and 1.7% of the entire area of the mainland respectively [28]. It is obvious that the statistical percentages are less than the corresponding P_t. The percentages of areas with positive intensity difference three or more are 0.14% and 0.10% for the two maps, while the earthquake death toll mainly, more than 95%, caused in these areas by large shocks. The main point from the test for improvement of PSHA is to recognize potential source areas of the future large earthquakes near big cities, especially those with magnitude more than 7.

8. Conclusion

From discussion on the fundamentals of SHA, strong earthquake may occur at a place where a destructive shock occurred with the similar magnitude, and strong earthquake may occur at a place where the tectonic condition is similar with another place where a destructive shock occurred, the state-of-the-art is reviewed as that PSHA hazard maps have done well at predicting the shaking from future earthquake in some cases, while they have done poorly in other cases. The authors concluded from formula derivation that Poisson distribution is derived just based on the independence assumption and the fact that the occurrence of destructive earthquake is rare. For estimation of seismicity parameters, two-rank scheme of potential source areas can manage data as many as possible, while it is quite difficult to evaluate the weighting factors from a comprehensive understand of the earthquake occurrence from seismic, tectonic and crust dynamic evidences. A logical approach to add up contributions of all these information is still requiring further study. The uncertainty correction procedure can combine the uncertainty of the attenuation into inherent uncertainty of ground motion, and can be improved with any new findings, such as the mean of random error of attenuation relations changing with acceleration. PSHA map could be tested from the observed intensity after they published without any subjective judgment and/or personal bias. The happened intensity exceeding that on the maps may be 2.67% and 1.7% from the case of two PSHA maps of China respectively, less than the corresponding 4.5% and 2.3% in 22 and 11 years from 10% in 50 years. The earthquake death toll mainly, more than 95%, caused in the areas with intensity seriously underestimated by large shocks. The main point to improve PSHA is suggested to recognize potential source areas of the future large earthquakes near big cities, especially those with magnitude more than 7.

Finally, the authors would like to emphasize that we are all on the way to a confident SHA, we have some data and know something on occurrence of strong earthquake, but cannot really predict it in a strict scientific way even for long-term, that fact is the reason of a P in PSHA. Therefore, hazard map either by PSHA or DSHA should be presented to the researchers, public and policy makers with a open detail technical report, for clear understanding of the uncertainties in it.

9. Acknowledgements

This work was financially supported by National Nature Science Foundation of China (51178151, 51178435 and 51478443), and International Science & Technology Cooperation Program of China (2011DFA21460).

10. References

[1] Castan H, Lomnitz C (2002): PSHA: is it science? *Engineering Geology*, **66**, 315-317.

[2] Stein S, R Geller, M Liu (2011): Bad assumptions or bad luck: Why earthquake hazard maps need objective testing. *Seismological Research Letters*, **82**(5), 623-626.

[3] Kossobokov V, Nekrasova A (2012): Global Seismic Hazard Assessment Program Maps are erroneous. *Seismic Instruments*, **48**(2), 162-170.

[4] Klügel J (2007): Error inflation in probabilistic seismic hazard analysis. *Engineering Geology*, **90**, 186-192.

[5] Wang Z, James C (2013): A critique of probabilistic versus deterministic seismic hazard analysis with special reference to the New Madrid seismic zone. *Geological Society of America*, Special Papers; **493**, 259-275.

[6] Kawazumi S (1951): Regional Distribution of Earthquake Hazard. *Journal of the Architectural Institute of Japan*, **77**, 3-9.

[7] Milne WG, Davenport AG (1969): Distribution of earthquake risk in Canada. *Bulletin of the Seismological Society of America*, **59**(2), 729-755.

[8] Gubin F (1962): *On issues of seismic zoning by seismo-tectonic procedure*. Science Press, Chinese version.

[9] Lee S (1957): The map of seismicity of China. *Acta Geophysica Sinica*, **6**(2), 127-158.

[10] Cornell CA (1968): Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, **58**, 1583-1606.

[11] Tao X, Tao Z, Shi L (2014): Scenario earthquake-Link of probabilistic seismic hazard assessment and deterministic seismic hazard assessment. *Earthquake Engineering and Engineering Dynamics*, **34**(4), 101-109 (in Chinese).

[12] Der KA, Ang AH (1977): A fault rupture model for seismic risk analysis. *Bulletin of the Seismological Society of America*, **67**(4), 1173-1194.

[13] Tao X (1982): An ellipsoid attenuation model. *Open file report*, Institute of Engineering Mechanics, Academic Sinica, No. 82-057 (in Chinese).

[14] Earthquake Research Committee (2000): *Long-term evaluation of the Miyagi- ken-Oki Earthquake*. (in Japanese)

[15] Earthquake Research Committee (2009): *Partial revision of long-term evaluation of seismic activity from Sanriku-Oki to Boso-Oki*. (in Japanese)

[16] Tao Z, Tao X (2012): A review on long-term evaluation of occurrence probability for subduction-zone earthquakes in Eastern Japan. *Applied Mechanics and Materials*, **166-169**, 2190-2196.

[17] Liu H (1987): On the seismic zoning map of China. *Proceedings of international seminar on seismic zonation*, Guangzhou, China. 35-39.

[18] Huang W, Li W, Cao X (1994): Part two of study on completeness of earthquake data in China continent—Regional distribution images of beginning years with complete earthquake data. *Acta Seismological Sinica*, **16**(4), 423-432.

[19] Swafford L, Stein S (2007): Limitations of the short earthquake record for seismicity and seismic hazard studies. *In Continental Intraplate Earthquakes*, Special Paper **425**, 49-58.

[20] Bommer JJ, Abrahamson NA (2006): Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bulletin of the Seismological Society of America*, **96**(6), 1967-1977.

[21] Strasser FO, Abrahamson NA, Bommer JJ (2009): Sigma: Issues, insights, and challenges. *Seismological Research*

Letters, **80**(1), 40-56.

[22] Esteva L (1969): Seismic risk and seismic design decisions. *Proceedings of MIT Symposium on Seismic Design of Nuclear Power Plants*, Cambridge, MA, US.

[23] Douglas J (2003): Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews*, **61**(1), 43-104.

[24] Campbell KW, Bozorgnia Y (2007): Campbell-Bozorgnia NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters. PEER *Report*

[25] Chiou BS, Youngs RR (2008): NGA model for average horizontal component of peak ground motion and response spectra. PEER *Report*

[26] Idriss IM (2007): Empirical model for estimating the average horizontal values of pseudo-absolute spectral accelerations generated by crustal earthquakes. *Interim Report* Issued for USGS Review

[27] Wang C, Tao X, Zharkova A, Zhao K, Jiang W (2014): A suggestion on uncertainty correction of ground motion attenuation in probabilistic seismic hazard assessment. *Applied Mechanics and Materials*, **580-583**, 1662-1666.

[28] Wang L, Tao X, Jiang W (2014): Preliminary inspection of Seismic Zoning Maps in China. *China Earthquake Engineering Journal*, **36**(4):1070-1076.