Probability-Based PGA Estimations Using Double-Lognormal Distribution: Example of the Nuclear Power Plant in Taiwan

Jui-Pin Wang
Assistant Professor, Dept. Civil & Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong. E-mail: jpwang@ust.hk.

Hamed Taheri
Research Student, Dept. Civil & Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong. E-mail: hamedsamira@ust.hk.

Thomas Chin-Tung Cheng
Ph.D, Sinotech Engineering Consultants, Inc., Taipei, Taiwan. E-mail: ctcheng@sinotech.org.tw

Keywords: double-lognormal distribution, seismic hazard assessment, nuclear power plant

ABSTRACT: Despite a variety of models available in earthquake forecasts and seismic hazard mitigations, earthquakes can not be perfectly predicted otherwise the recent earthquake-induced disasters can be prevented or mitigated. This study presents an approach in seismic hazard assessments. The site-specific seismic hazard is portrayed by the relationship between an estimated earthquake motion and its annual exceedance probability. The approach is characterized with the use of double-lognormal distribution. Its use in the simulation of the earthquake motion was verified by statistical testing. The approach was demonstrated and used to evaluate the seismic hazard at the 4th nuclear power plant in Taiwan. The result shows that the peak ground acceleration is estimated at 0.53 g considering the annual exceedance probability equal to 1%.

1. INTRODUCTION

Taiwan is well known to its unique tectonic setting that has resulted in a high seismicity around the region. A few catastrophic earthquakes, the Chi-Chi earthquake in 1999 for instance, have pounded the island and caused a lot of damages. It is understood that it is not a matter of "yes-or-no" about the next catastrophic earthquake striking Taiwan; instead, it is a matter of "when", "where" and "how large".

The local community of geosciences and civil engineering has spent lots of efforts on a variety of research regarding earthquake engineering and hazard mitigation. The Central Geological Survey, Taiwan, has investigated the active faults in Taiwan and the results are updated and published periodically (Lin *et al*., 2008, 2009). A few earthquake early warning systems (EEWS) have been developed for the region around Taiwan (e.g., Wu *et al*., 2003, 2005). In probabilistic and statistical analyses, Wang *et al.* studied the distribution of annual maximum earthquakes since 1900 and suggested a 5%-probability for a catastrophic earthquake (magnitude $\geq$ 7.5) occurring in Central Taiwan in next 50 years. However, none of the studies can yet be applicable to evaluate site-specific seismic hazard for civil engineering designs as the probabilistic seismic hazard analysis (PSHA) proposed in the late 60s' (Cornell 1968). A number of studies on regional seismic hazard evaluations have been conducted by using PSHA (e.g., Frankel *et al*., 1996; Cheng *et al*., 2007; Nakajima *et al*., 2008; Roshan and Basu, 2010). In the last decade, the PSHA has been a standard approach for seismic hazard assessments at the nuclear power plant sites (IAEA, 2002; U.S. Nuclear Regulatory Commission, 2007).

Here we present an approach that is characterized with the use of double-lognormal distribution in seismic hazard assessments. The site-specific, probability-based seismic hazard is portrayed by the relationship between a given earthquake motion and its exceedance probability. The methodology is detailed in this paper and the approach was demonstrated and used to evaluate the seismic hazard at the 4th nuclear power plant sites in Taiwan.

2. METHODOLOGY

2.1 Ground motion models

Ground motion models are used to estimate the ground motion (Y) at a site. Their general expression is:

$$\ln Y = f(\theta_1,...,\theta_n) = f(\mathbf{\Omega}) \;;\; \sigma_{\ln Y} = \sigma^* \quad (1)$$

where $\theta_1,\ldots,\theta_n$ denote earthquake variables, e.g., magnitude, distance, etc; $\mathbf{\Omega}$ denotes the set of the earthquake variables $\theta_1,\ldots,\theta_n$. Note that the number of earthquake variables depends on the formation of a ground motion model. The model uncertainty is presented by a constant (σ^*). An important feature of the model is that the logarithm of Y follows the normal distribution; alternatively, Y follows the lognormal distribution. The mean and standard deviation (S.D.) of $\ln Y$ are $f(\mathbf{\Omega})$ and σ^*, respectively.

Since Y is a random variable following the lognormal distribution, its 50th and 84th percentiles can be presented as Eq.2 by extending Eq.1.

$$Y_{50} = \ln Y_{50} = f(\mathbf{\Omega}) \quad (2.1)$$

$$Y_{84} = \ln Y_{84} = f(\mathbf{\Omega}) + \sigma^* \quad (2.2)$$

2.2 Semi-observed ground motion

A series of ground motions at a given site can be "back-calculated" provided an earthquake catalog is available. Note that ground motions are not directly measured during earthquake occurrences. They are the "best estimates" based on observed earthquakes along with the use of proper ground motion models. As a result, this study refers them as semi-observed ground motions.

2.3 Exceedance probability based on a double-lognormal distribution

The proposed approach was developed based on the characteristics of the variable, semi-observed ground motion Y. If it can be modelled by a probability distribution, the exceedance probability associated a given y^* can be estimated. Provided it follows the double-lognormal distribution, the exceedance probability becomes:

$$\begin{aligned} P(Y > y^* \mid \mu_d, \sigma_d) &= P(\ln Y > \ln y^* \mid \mu_d, \sigma_d) \\ &= 1 - P(\ln Y \le \ln y^* \mid \mu_d, \sigma_d) \\ &= 1 - \Phi\left(\frac{\ln(\ln(y^*)) - \mu_d}{\sigma_d}\right) \end{aligned} \quad (3)$$

where μ_d and σ_d denote the mean and S.D. of the double-logarithm of Y, i.e., $\ln(\ln(Y))$; Φ denotes the cumulative probability function of a standard normal variate.

2.4 Probability-based seismic hazard

In PSHA, seismic hazard is originally presented by the annual rate of exceedance (λ_{y^*}) associated with a given motion y^*. Note that the rate-based seismic hazard is not a probability-based quantity since it is allowed to exceed 100%. In this study, the seismic hazard is probability-based, portraying the probability that at least one earthquake motion exceeds a given motion within the duration of interest. This probability is referred as exceedance probability in this study. Two schemes were developed to evaluate this seismic hazard.

The first scheme is to consider the rate of earthquake occurrence a constant. Provided the rate is v within a given period, the exceedance probability can be estimated by extending Eq.3. Since the exceedance probability is equal to 100% minus the probability that not a single motion exceeds y^*, it becomes:

$$P(Y > y^* \mid \mu_d, \sigma_d, v) = 1 - \left(P(Y \le y^* \mid \mu_d, \sigma_d)\right)^v \quad (4.1)$$

where

$$P(Y \le y^* \mid \mu_d, \sigma_d) = \Phi\left(\frac{\ln(\ln(y^*)) - \mu_d}{\sigma_d}\right) \quad (4.2)$$

In contrast, the second scheme considers the rate a random variable. The variable is assumed to follow the Poisson distribution. Note that the assumption used in modelling earthquake occurrences through time is adopted in existing seismic hazard analyses, i.e., PSHA. Eq.5 shows the probability mass function of the Poisson distribution.

$$P(N = n \mid v) = \frac{v^n \times e^{-v}}{n!} \quad (5)$$

where N denotes a discrete random variable. Combining Eqs. 4.1 and 5, the exceedance probability becomes:

$$\begin{aligned} &P(Y > y^* \mid \mu_d, \sigma_d, v) \\ &\quad = 1 - \sum_{n=0}^{\infty} \frac{v^n \times e^{-v}}{n!} \times P(Y \le y^* \mid \mu_d, \sigma_d)^n \\ &\quad = 1 - e^{-v}\left[1 + \frac{v \times P(Y \le y^* \mid \mu_d, \sigma_d)}{1!} + \right. \\ &\qquad \left. + \cdots + \frac{v^{\infty} \times P(Y \le y^* \mid \mu_d, \sigma_d)^{\infty}}{\infty!}\right] \end{aligned} \quad (6)$$

The term in the bracket is the Taylor expansion of $e^{v \times P(Y \le y^* \mid \mu_d, \sigma_d)}$. Therefore, Eq. 6 becomes:

$$\begin{aligned} P(Y > y^* \mid \mu_d, \sigma_d, v) &= 1 - e^{-v} e^{v \times P(Y \le y^* \mid \mu_d, \sigma_d)} \\ &= 1 - e^{v \times (P(Y \le y^* \mid \mu_d, \sigma_d) - 1)} \end{aligned} \quad (7)$$

where the probability term has been shown in Eq.4.2.

2.5 Goodness-of-fit testing: K-S test

The probability-based seismic hazard by Eqs.4 and 7 is based on the presumption that Y follows the double-lognormal distribution. As a result, statistical goodness-of-fit needs to be performed on the assumption. This study uses the Kolmogorov-Smirnov (K-S) test in the goodness-of-fit testing. The K-S test is to compares the maximum difference between observational and theoretical probabilities. When the maximum difference is larger than the critical value in testing, the model is considered not suitable in modelling the variable. Eq.8 shows the mapping of the observational probability in the K-S test.

$$S_n(x) = \begin{cases} 0 & x < x_1 \\ k/n & x_k \le x < x_{k+1} \\ 1 & x \ge x_n \end{cases} \quad (8)$$

where n is the number of observation; x_k denotes the k^{th} observation in an ascending order.

3. SEISMIC HAZARD AT THE 4TH NPP IN TAIWAN

3.1 Inputs

Fig. 1 shows the seismicity (more than 50,000 earthquakes) around Taiwan since 1900. The 4th nuclear power plan is located in the north-eastern region of Taiwan. Table 1 summarizes the ground motion models that were used in the following analysis. Since the magnitude units are different in the earthquake catalog (local magnitude M_L) and ground motion model (moment magnitude M_w), a relationship (Eq. 9) is needed to convert from one to the other.

$$M_L = 4.533 \times \ln(M_w) - 2.091 \quad (9)$$

The inputs used in the analysis were developed and used in earthquake-related studies focusing on the region around Taiwan (Wu *et al*., 2002; Cheng *et al*., 2007; Wang *et al*., 2011).

Table 1. Ground motion models suitable for the region of Taiwan

Description	PGA Attenuation Relationship	σ*
Hanging wall, Rock, Crustal	ln y = -3.25 + 1.075 M_W -1.723 ln(R + 0.156 exp(0.62391 M_W))	0.577
Hanging wall, Soil, Crustal	ln y = -3.25 + 1.075 M_W -1.723 ln(R + 0.156 exp(0.62391 M_W))	0.555
Foot wall, Rock, Crustal	ln y = -3.25 + 1.075 M_W -1.723 ln(R + 0.156 exp(0.62391 M_W))	0.583
Foot wall, Soil, Crustal	ln y = -3.25 + 1.075 M_W -1.723 ln(R + 0.156 exp(0.62391 M_W))	0.554

3.2 Distance and Magnitude Thresholds

Distance and magnitude thresholds (D_0 and M_0) are involved in obtaining the series of semi-observed ground motions. As shown in Fig. 1, a number of earthquakes have occurred since 1900 but only a small percentage among them was large earthquakes (e.g., magnitude ≥ 6.5) which possibly cause damages in civil infrastructures and facilities. In addition, a large earthquake is also very unlikely to cause damages in far fields. Therefore, the distance and magnitude thresholds were used to filter out those earthquakes that have a negligible probability causing structural damages. On the other hands, those with magnitude greater than M_0 and source-to-site distance less than D_0 are referred as "featured" earthquakes in this study. The corresponding series of semi-observed ground motions are contributed by the series of featured earthquakes.

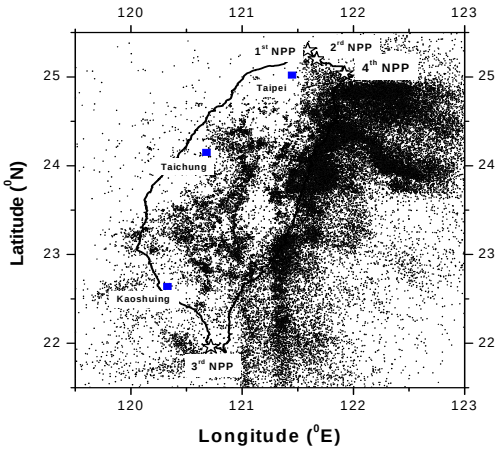

Fig. 1 Seismicity around Taiwan since 1900

3.3 Semi-observed ground motion near the 4th NPP

Fig. 2(a) shows the 123 featured earthquakes considering M_0 and D_0 of 6.0-M_w and 200-km, respectively, near the 4th NPP since 1900. The annual mean occurrence rate is equal to 1.12 accordingly. Fig. 2(b) shows the variation of the 50th percentile of semi-observed ground motion Y_{50}. The series of Y_{50} presents an asymmetrical distribution against its mean. As a result, Y_{50} is unlikely to follow the normal distribution that is symmetrical. Figs. 2(c) and 2(d) shows the distributions of ln(Y_{50}) and ln(ln(Y_{50})). It was found that the level of symmetry was improved, which indicates that ln(ln(Y_{50})) possibly follows the normal distribution associated with its mean and S.D. of 0.98 and 0.31, respectively.

The K-S test was performed at a significance level of 5% to examine whether ln(ln(Y_{50})) follows the normal distribution. Fig. 3 shows the observational and theoretical distributions in cumulative probability. The maximum difference is 0.08 less than the critical value

of 0.12. It indicates that ln(ln(Y_{50})) can be modelled by the normal distribution. In other words, Y_{50} follows the double-lognormal distribution.

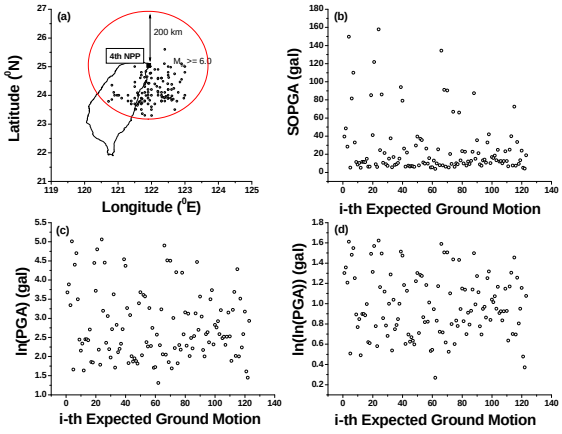

Fig. 2 Distributions of 123 semi-observed ground motion near the 4th NPP

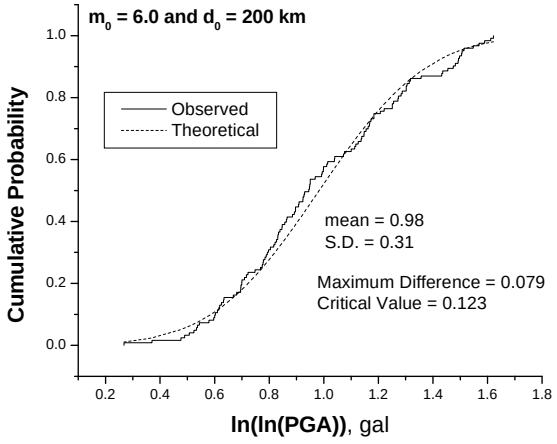

Fig. 3 Cumulative frequency for the K-S test of ln(ln(Y_{50})) near the 4th NPP

3.4 Probability-based seismic hazard

After the ground motion was verified to follow the double-lognormal distribution, the probability-based seismic hazard at the 4th NPP site can be estimated by using Eqs.4 and 7. Fig. 4 shows the hazard based on the two schemes. The peak ground accelerations are 0.26 g at a given annual exceedance probability equal to 1%. The difference in the estimations between two schemes is negligible.

3.5 Effect of distance and magnitude thresholds

The relationship shown in Fig. 4 is under a specific magnitude and distance thresholds for the 4th NPP sites. It will change when different thresholds are used. More importantly, the double-lognormal distribution could be no longer suitable in modelling Y_{50}. Therefore, additional five sets of thresholds, which are (6.0, 150 km), (6.0, 250 km), (5.5, 200 km), (5.5, 150 km) and (5.5, 100 km) were adopted and the analyses were repeated. Fig.5 shows the mean annual rates, means of ln(ln(Y_{50})), coefficients of variation (COV) of ln(ln(Y_{50})), maximum differences, and critical values for each threshold set. The maximum differences are less than the critical values in the K-S test for each set (Fig. 5(a)). This verified that the ground motion at the 4th NPP site follows the double-lognormal distribution in these conditions.

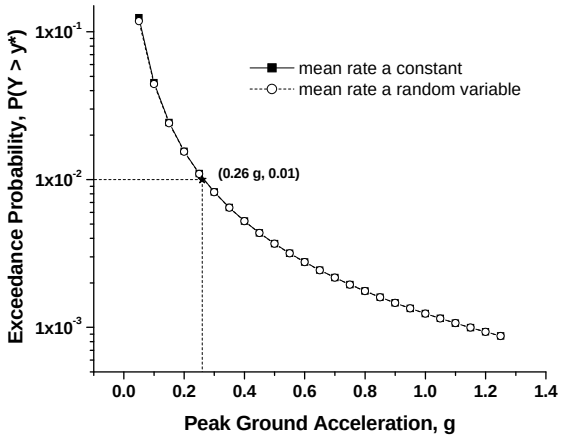

Fig. 4 Seismic hazard curve based on the threshold set of (6.0-M_w, 200-km)

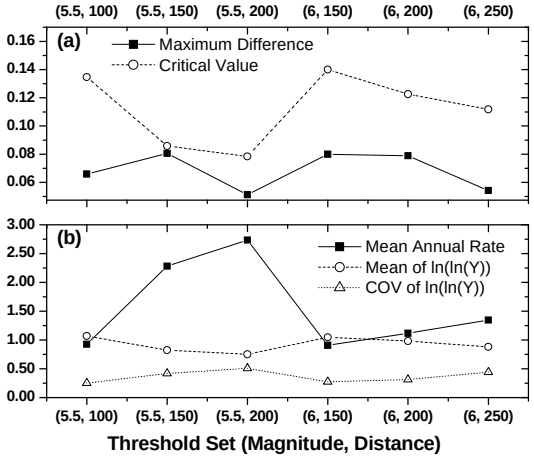

Fig. 5 Statistics of ln(ln(Y_{50})) for each threshold set

Fig. 6 shows seismic hazards by the proposed approach associated with the statistics shown in Fig. 5(b). In both results, the highest and lowest seismic hazards are governed by using the threshold sets of

(6.0-M_w, 250-km) and (5.5-M_w, 100-km), respectively. The mean value of ln(ln(Y_{50})) considering (6.0-M_w, 250-km) is in fact lower than that considering (5.5-M_w, 100-km). However, its COV is relatively high. Therefore, the probability distribution is characterized with "long tails" due to its high variability, leading to a high probability associated with extreme values.

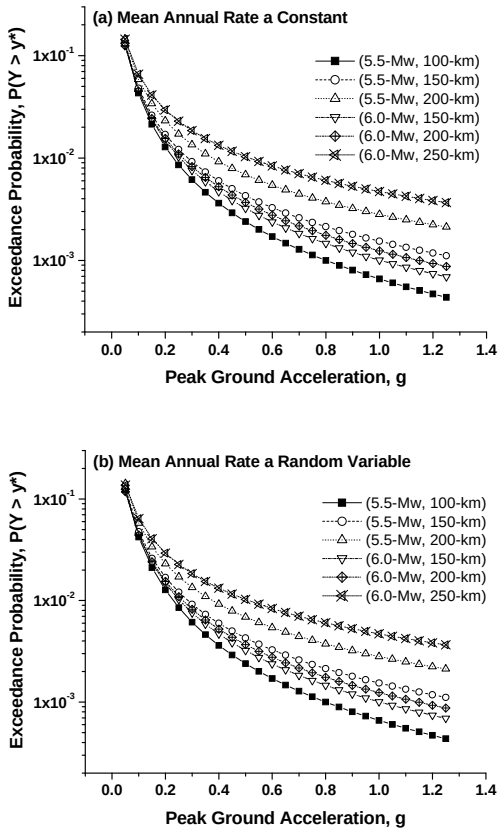

Fig. 6 Seismic hazard curves based on ln(ln(Y_{50})) under different threshold sets

3.6 Effect of ground motion model variability

The results that have been shown are based on the use of the variable ln(ln(Y_{50})). Its expression, as shown in Eq. 2.1, does not consider the uncertainty of ground motion models (σ*). In order to consider the effect, the analysis was repeated by using the variable ln(ln(Y_{84})) which has been expressed in Eq.2.2.

Fig. 7 shows the statistics of ln(ln(Y_{84})) for each threshold set. When (5.5-M_w, 150-km) was used, the double-lognormal distribution is not suitable in modelling ln(ln(Y_{84})) although the critical value is barely greater than the maximum difference. Compared with ln(ln(Y_{50})), the mean value of ln(ln(Y_{84})) increases as expected but the COV remains a comparable level. Fig. 8 shows the seismic hazard associated with the statistics of ln(ln(Y_{84})) shown in Fig. 7(b). The pattern is similar to that shown in Fig. 6. The highest and lowest seismic hazards are still governed by the thresholds of (6.0-M_w, 250-km) and (5.5-M_w, 100-km), respectively.

3.7 Representative seismic hazard curves

Figs. 6 and 8 have shown the seismic hazard curves under a variety of situations. The differences were found relatively significant. Two schemes were used to estimate the representative hazard curves incorporating the various estimations. One uses the average and the other uses the envelope (maximum) among the twelve curves. The latter adds more conservatism in seismic hazard estimations. Note that if ground motion can not be satisfactorily modelled by the double-lognormal distribution under a specific threshold set, the estimation is not considered in the determination of the "averaged" representative curve. Fig. 9 shows the representative curves. The envelope curve is governed by the use of ln(ln(Y_{84})) and (6.0-M_w, 250-km). Given an annual exceedance probability equal to 1%, the corresponding PGAs are 0.53 g (envelope) and 0.31 g (average) estimated by the two schemes.

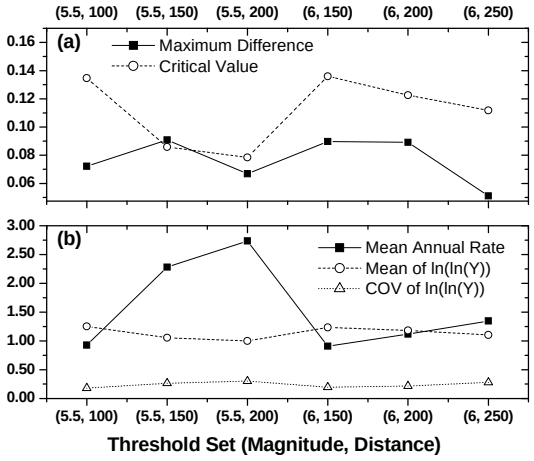

Fig. 7 Statistics of ln(ln(Y_{84})) for each threshold set

4. CONCLUSIONS AND DISCUSSIONS

This paper presents an approach to evaluate seismic hazard portrayed by the relationship between a given motion and its exceedance probability. The approach is characterized with a probabilistic framework associated with the use of the double-lognormal distribution. The relationship was verified by the series of ground motions near the 4th NPP in Taiwan back-calculated based on the regional seismicity since 1900. It shows that the seismic hazard estimations are 0.53 g (envelope scheme) and 0.31 g (average scheme) considering

annual exceedance probability equal to 1%. Using the envelope scheme, the estimations are governed by the use of threshold set (6.0-M_w, 250-km).

The seismic hazard estimations for the 4th NPP site shown in this paper are based on the double-lognormal distribution that was verified satisfactorily in modelling either 50th or 84th percentile ground motions. If the approach is used to evaluate the seismic hazard at different sites, the verification on the suitability of the double-lognormal distribution must be performed at the first place. The hazard estimations are problematic when the assumption is not verified or it is proved unsatisfactory.

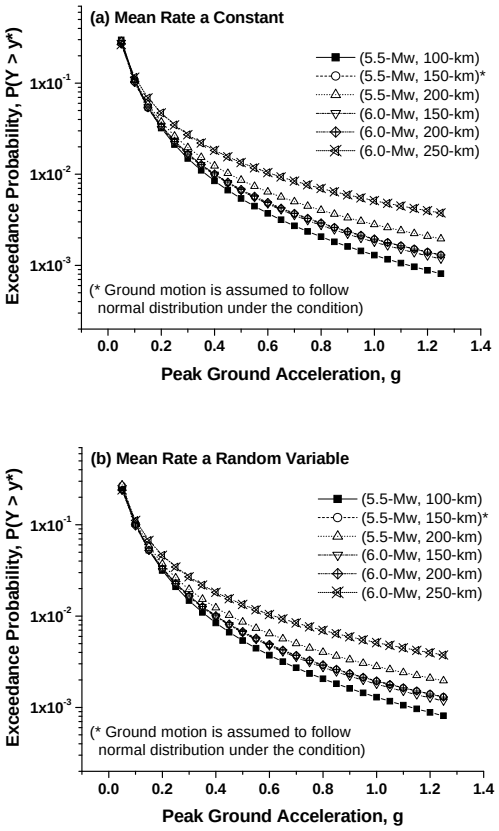

Fig. 8 Seismic hazard curves based on ln(ln(Y_{84})) under different situations

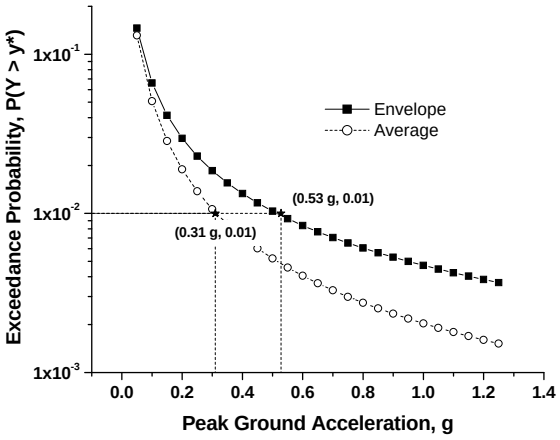

Fig. 9 Representative seismic hazard curves at the 4th NPP in Taiwan

6. REFERENCES

Cheng, C.T., Chiou, S.J., Lee, C.T., Tsai, Y.B. (2007). Study on probabilistic seismic hazard maps of Taiwan after Chi-Chi earthquake. *Journal of GeoEngineering,* 2, pp 19-28.

Cornell C.A. (1968) Engineering seismic risk analysis. *Bull. Seism. Soc. Am.* 58, pp 1583-1606.

Frankel, A., Mueller, C., Barnhard, T., Perkins, D., Leyendecker, E. V., Dickman, N., Hanson, S., M. Hopper (1996). National seismic-hazard map: documentation June 1996. *USGS Open-File Report*, pp 96-532.

International Atomic Energy Agency (2002). Evaluation of seismic hazards for nuclear power plant. *Safety Guide NS-G-3.3*.

Lin, C.W., Lu, S.T., Shih, T.S., Lin, W.H., Liu, Y.C. and Chen, P.T. (2008). Active Faults of Central Taiwan. *Special Publication of Central Geological Survey.* 21: pp 148. (In Chinese with English abstract)

Lin, C.W., Chen, W.S., Liu, Y.C., Chen, P.T. (2009) Active Faults of Eastern and Southern Taiwan. *Special Publication of Central Geological Survey.* 23: 178. (In Chinese with English abstract)

Nakajima, M., Choi, I.K., Ohtori, Y., Choun, Y.S., (2008) Evaluation of seismic hazard curves and scenario earthquakes for Korean sites based on probabilistic seismic hazard analysis. *Nuclear Engineering and Design* 237, pp 277-288.

Roshan, A.D., Basu, P.C. (2010). Application of PSHA in low seismic region: A case study in NPP site in peninsular India. *Nuclear Engineering and Design* 240, pp 3443-3454.

U.S. Nuclear Regulatory Commission (2007). A performance-based approach to define the site-specific earthquake ground motion. *Regulatory*

Guide 1.208.

Wang, J.P., Chan, C.H., Wu, Y.M. (2011), The distribution of annual maximum earthquake magnitude around Taiwan and its application in estimation of catastrophic earthquake recurrence probability. *Natural Hazard*, DOI: 10.1007/s11069-011-9776-x (in press)

Wu, Y.M., and Teng, T.L. (2002). A virtual sub-network approach to earthquake early warning. *Bull. Seism. Soc. Am.*, 92, pp 2008-2018.

Wu, Y.M., Teng, T.L., Shin, T.C. and Hsiao, N.C. (2003). Relationship between peak ground acceleration, peak ground velocity, and intensity in Taiwan. *Bull. Seism. Soc. Am*. 93, pp 386-396

Wu, Y.M. and Kanamori, H. (2005). Experiment on an onsite early warning method for the Taiwan early warning system. *Bull. Seism. Soc. Am*. 95, pp 347-353.