Use of the T-distribution to Construct Seismic Hazard Curves for Seismic Probabilistic Safety Assessments

Eric Yee[1]

[1] *KEPCO International Nuclear Graduate School, 658-91, Haemajiro, Seosaeng-myeon, Ulju-gun, Ulsan 45014, and eric.yee@kings.ac.kr*

Seismic probabilistic safety assessments are used to help understand the impact potential seismic events can have on the operation of a nuclear power plant. An important component to seismic probabilistic safety assessment is the seismic hazard curve which shows the frequency of seismic events. However, these hazard curves are estimated assuming a normal distribution of the seismic events. This may not be a strong assumption given the number of recorded events at each source-to-site distance. Using a normal distribution makes the calculations significantly easier but may underestimate or overestimate the more rare events, which is of concern to nuclear power plants. This paper shows a preliminary exploration into the effect of using a distribution that perhaps more represents the distribution of events, such as the t-distribution to describe data. The integration of a probability distribution with potentially larger tails basically pushes the hazard curves outward, suggesting a different range of frequencies for use in seismic probabilistic safety assessments. Therefore the use of a more realistic distribution results in an increase in the frequency calculations suggesting rare events are less rare than thought in terms of seismic probabilistic safety assessment. However, the opposite was observed with ground motion prediction equation considered.

I. INTRODUCTION

A common technique to evaluate the exposure a nuclear power plant, or any other important piece of infrastructure, to earthquakes is to perform a probabilistic safety assessment for seismic events. The state-of-practice takes on an interdisciplinary approach, with multiple components that designers and engineers must consider. One integral component to seismic probabilistic safety assessments is the seismic hazard curve. This curve basically plots an intensity measure, such as peak ground acceleration (PGA) or peak ground velocity, to an annual probability of exceedance (APE). Some hazard curves use a return period in place of an APE, but the name "return period" has caused some confusion to practitioners over the years. Another component is a system, structure, or component fragility curve. The fragility curve maps the conditional probability of a system, structure, or component failure due to a certain intensity measure. For the same intensity measure, the seismic hazard curve and the fragility curve are multiplied to produce a failure frequency, or failure probability curve for a specific system, structure, or component. The area under this failure probability curve would be considered the failure rate. These values are then compared to satisfy a design or regulatory minimum threshold.

Seismic hazard curves can be calculated deterministically or probabilistically. Probabilistic seismic hazard analysis is used in constructing the seismic hazard curve. Probabilistic seismic hazard analyses essentially have four steps. First, a characterization of the site for potential seismic sources is performed, which involves magnitude and source-to-site distance identification in a probabilistic manner. The second step is the establishment of a magnitude recurrence relationship, which addresses the temporal issues in the analyses. The third step is to prepare a ground motion prediction equation (GMPE). The GMPE takes in earthquake source characteristics and source-to-site distances to estimate a variety of intensity measure parameters. Modern GMPEs usually treat traditional intensity measures as log-normally distributed. The fourth step uses the GMPE and temporal relationships to construct the seismic hazard curve.

Common criticism to GMPEs is the lack of data for certain conditions, such as earthquake magnitude (M) and site-to-source distance (R). As an example, Fig. 1 shows a sample database of recorded seismic events up to 2008. The figure shows a lack of data for distances less than about 10 to 20 km, and for low magnitudes generally. This presents an issue where it is difficult for the engineer to confine or model earthquake effects for those conditions. It is generally accepted that the behavior observed at moderate distances where there is a good amount of data, should be applicable to close and far distances. However, earthquake engineers understand that observed behavior at close and far distances may not be properly represented by those at moderate distances. For example, issues such as radiation damping, surface waves, and ground nonlinearity make it difficult to address.

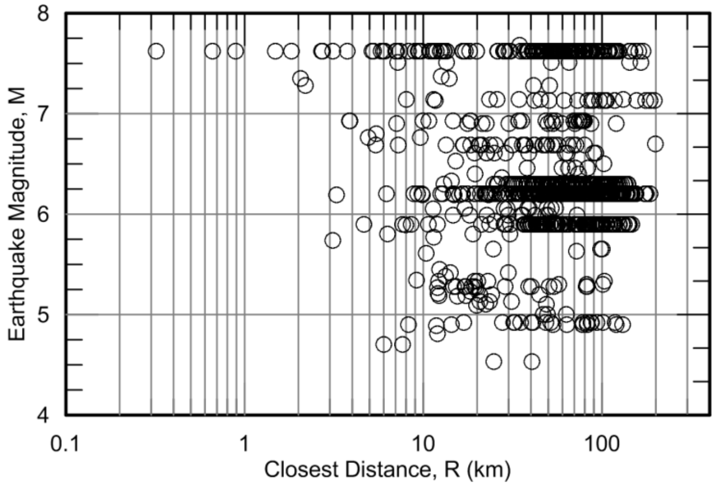

Fig. 1. Projected number of earthquakes used by Idriss[1] up to 2008.

This paper attempts to address this data issue by assigning a student's t-distribution in place of the traditional normal distribution to the natural logarithm of a sample intensity measure. This requires the deconstruction of a popular GMPE to evaluate under which conditions a t-distribution would be applicable. A sample scenario is then presented to showcase the effects, if any, a change in distribution would have on the results of the simple probabilistic safety assessment.

II. HAZARD CURVE CONSTRUCTION

II.A. De-construction of Selected Ground Motion Prediction Equation

In order to determine under which conditions a t-distribution might be applicable, a deconstruction of a GMPE would be needed. This paper selected the GMPE by Idriss[1] as this GMPE is: (1) easy and simple to use, (2) data used to construct the GMPE is readily available, and (3) the logic and data selection process is generally well explained since it is part of the larger Next Generation Attenuation program[2,3]. Following the descriptions by Idriss[1], a database was developed containing the earthquake events, recording stations, and PGA data. For simplicity, only PGA will be considered as the intensity measure for this study. These events were then compared to the Idriss GMPE[1] to calculate residuals, as shown in Fig. 2.

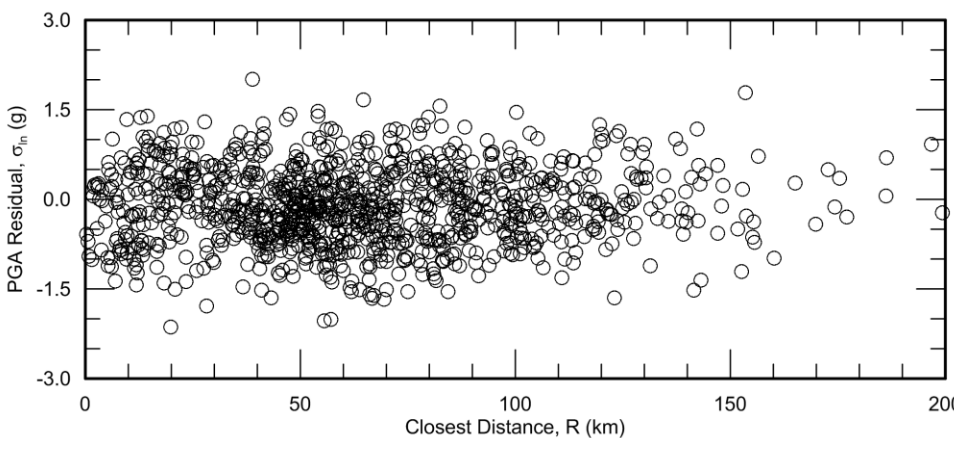

Fig. 2. Reverse-engineered residuals using the event and station descriptors by Idriss[1].

Fig. 3 plots the PGAs from seismic events recorded at distilled recording stations against the derived GMPE for normal and reverse fault mechanisms, where mechanism 0 is for normal and strike-slip events and mechanism 1 is for reverse faulting events. The plots show two sets of data, when $4.5 \leq M < 5$ and when $5 \leq M$ 5.5, which is similar to how Idriss[1] grouped his data set. The indicated earthquake magnitude is the average of all the data in their corresponding range. These plots show a lack of data at distances of less than about 10 km for these magnitude ranges. The lack of data for distances greater than 150 km is because the designer limited the GMPE to less than 200 km and that observation that attenuation eventually reduces the ground motions to negligible recordings.

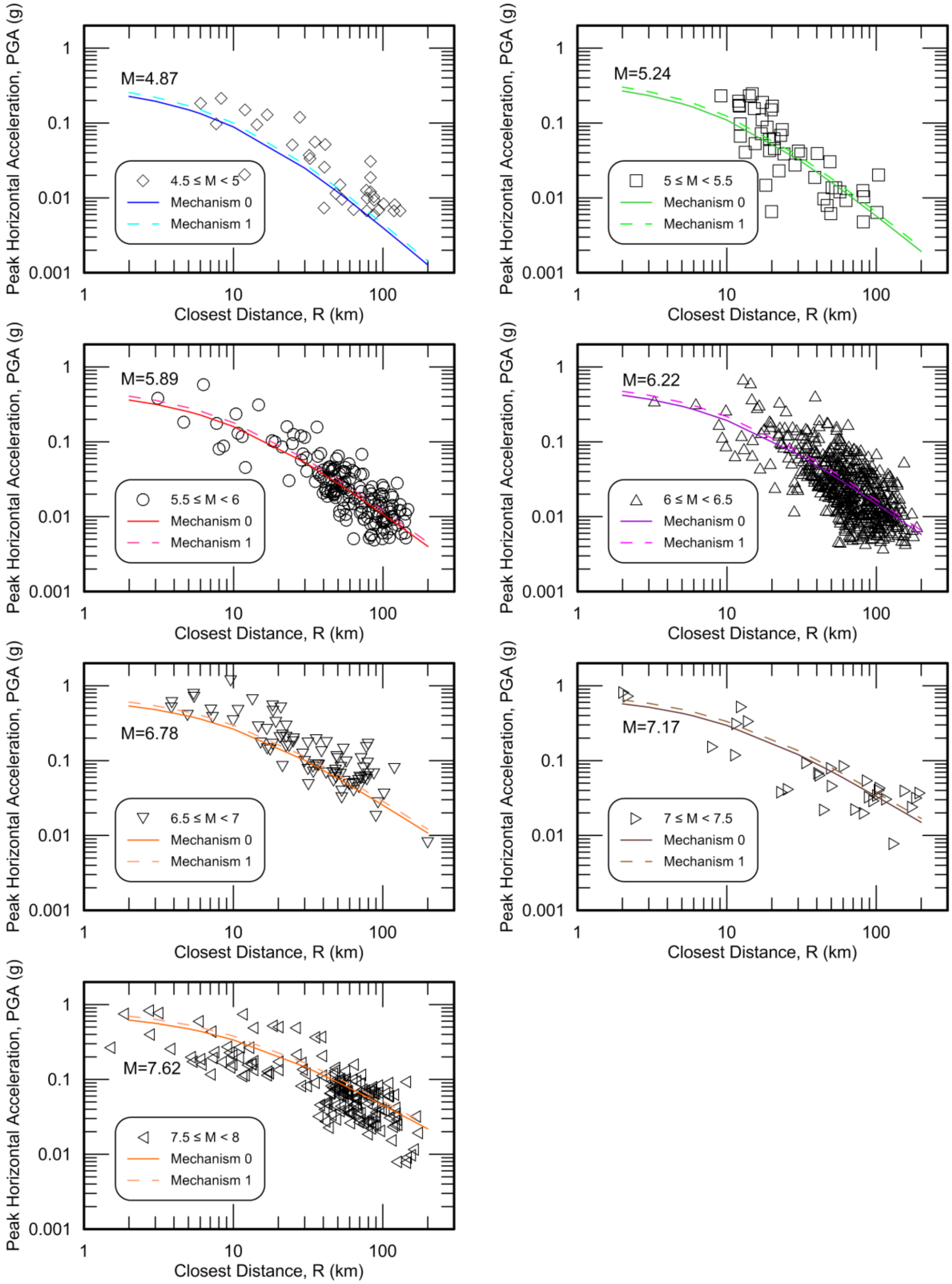

Fig. 3. Reverse-engineered data points and the median GMPE under different seismic source conditions for Idriss[1].

When compared to the results provided in Idriss[1], we were unable to reverse engineer a perfect match. Some stations were unable to be reasonably eliminated and some of the station descriptors listed in Idriss[1] were not identified in the updated flatfile provided by Pacific Earthquake Engineering Research Center (PEER). This resulted in relatively more data points, mostly from the Taiwan events. Fig. 2 basically shows more points than the figures provided in Idriss[1] as well as some points having differing residuals. It would seem that some of that data used by Idriss[1] were different than that provided in the updated flatfile by PEER.

II.B. Scenario

The scenario considered for this preliminary study is shown in Fig. 4. This shows a site with two earthquake sources defined as vertical linear faults. Fault A has a closest site-to-source distance of 10 km that can only produce earthquakes with magnitude 6.5 at a rate of 0.01 times a year. Fault B has a closest site-to-source distance of 20 km that can only produce earthquakes with magnitude 7.5 at a rate of 0.002 times a year. This is a common example scenario and is only used here to demonstrate the effects of change in GMPE characteristic.

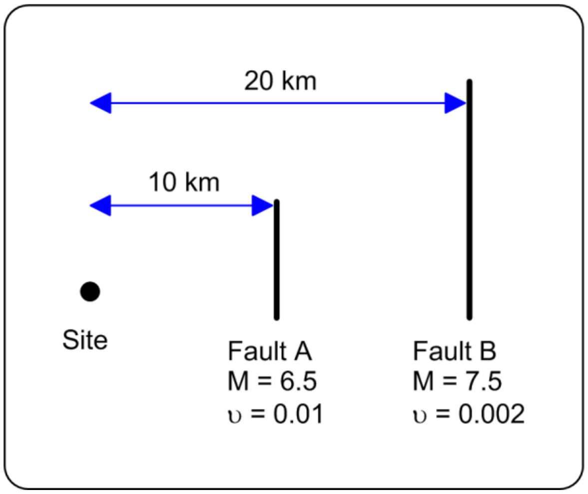

Fig. 4. Map view of example site with two earthquake sources.

For the given scenario, the hazard integral is calculated using the original Idriss[1] GMPE and is plotted in Fig. 5. Hazard curves are calculated from the hazard integral by specifying a target PGA and then calculating the probability of exceeding the aforementioned level of PGA. This is calculated by taking the area under the lognormal distribution, which is the underlying distribution of the PGA values, given distance and magnitude. The area is typically calculated by using the z-value. The sum of all the areas at each source-to-site distance represents the hazard level APE in the hazard curve. An example is shown in Fig. 6. As one can see, if the distribution changes, so should the hazard level. Additionally, a change in the distribution would also result in a change in the GMPE. Fig. 5 also plots the results of using the t-distribution in the Idriss[1] GMPE. The t-distribution was basically used by estimating a t-value, which is essentially the same as a z-value for most GMPEs, by subtracting the median GMPE value from the PGA level and then dividing by a standard deviation for the t-distribution. The standard deviation for the unknown t-distribution was calculated by Eq. (1) where x_i = data points, $\bar{x}$ = GMPE median, n = number of points. The limits are from 0.1 to 2 g because PGAs less than 0.1 g are of little interest while those above 2 g are slightly unrealistic.

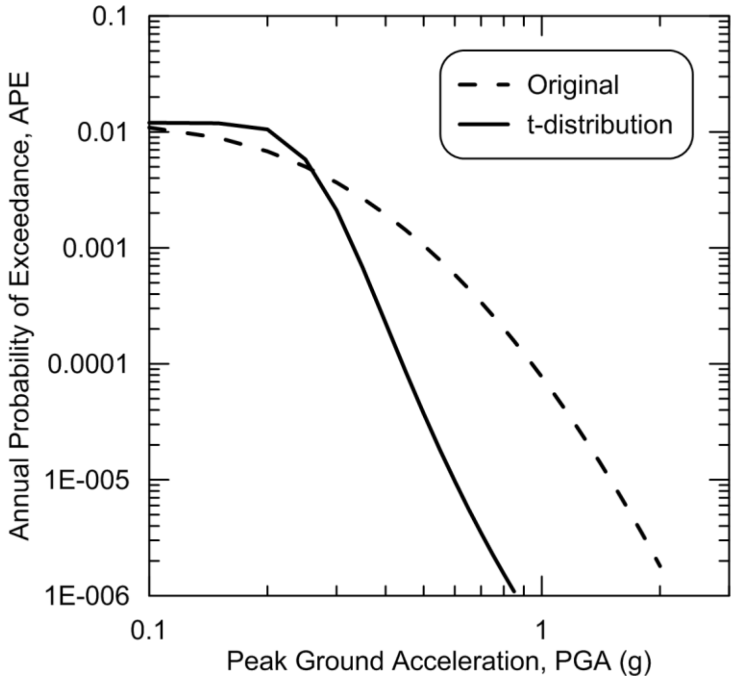

Fig. 5. Hazard curves.

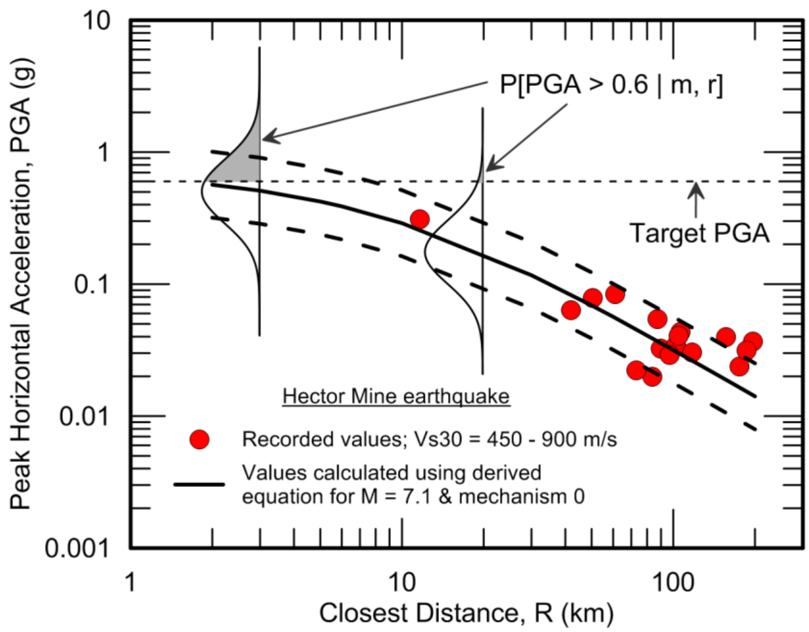

Fig. 6. Example hazard curve calculations using Idriss[1] data for Hector Mine event.

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n)(n-1)}} \qquad (1)$$

Fig. 5 shows that by changing to a student's t-distribution the hazard curve shows a reluctance to reach large accelerations. For accelerations less than about 0.25 g, the adjusted hazard curve appears to show an increase in hazard APE, while for accelerations greater than about 0.25 g, the adjusted hazard curve appears to show a decrease in hazard APE relative to the original Idriss[1] GMPE for the subject scenario. This strongly suggests that relatively strong seismic events are overestimated relative to the selected GMPE, which also suggests the t-distribution used most likely had smaller tails than the log-normal distribution estimated by the GMPE. Table 1 shows that the standard deviations for the log-normal distribution were relatively larger than the standard deviations for the t-distribution of the natural logarithm of the PGAs. For fault A, there were 9 data points while for fault B, there were 43 data points at distances equal to and less than the closest distance shown in Fig. 4. The results are slightly contrary to what was expected.

TABLE I. List of standard deviations for the earthquake events

Seismic source	Normal distribution $\sigma_{\ln}$	Student's t-distribution $\sigma_{\ln}$
Fault A	0.61	0.19
Fault B	0.53	0.12

II. FRAGILITY CURVE

For the scenario, we will only consider the generic fragility curve for a pressurizer. As real data is not readily available, the generic fragility curve is calculated using the procedure outlined by EPRI[4]. This procedure basically uses Eq. (2):

$$f' = \Phi\left[\frac{\ln\left(\frac{a}{A_m}\right) + \beta_U \Phi^{-1}(Q)}{\beta_R}\right] \tag{2}$$

where f' = conditional probability of failure, a = given peak ground acceleration level, A_m = median ground acceleration capacity, β_U = inherent randomness (aleatory uncertainty), β_R = unknown uncertainty (epistemic uncertainty), Q = subjective probability (confidence) that the conditional probability of failure, f, is less than f' for a peak ground acceleration a, Φ^{-1} = inverse of Gaussian cumulative distribution, and Φ = standard Gaussian cumulative distribution. For a pressurizer, the parameters are taken from NUREG/CR-6544[5] with A_m = 2.5 g, β_U = 0.40, and β_R = 0.30. Additionally, Q = 50% was chosen. Fig. 7 plots the results for an example pressurizer using the aforementioned parameters.

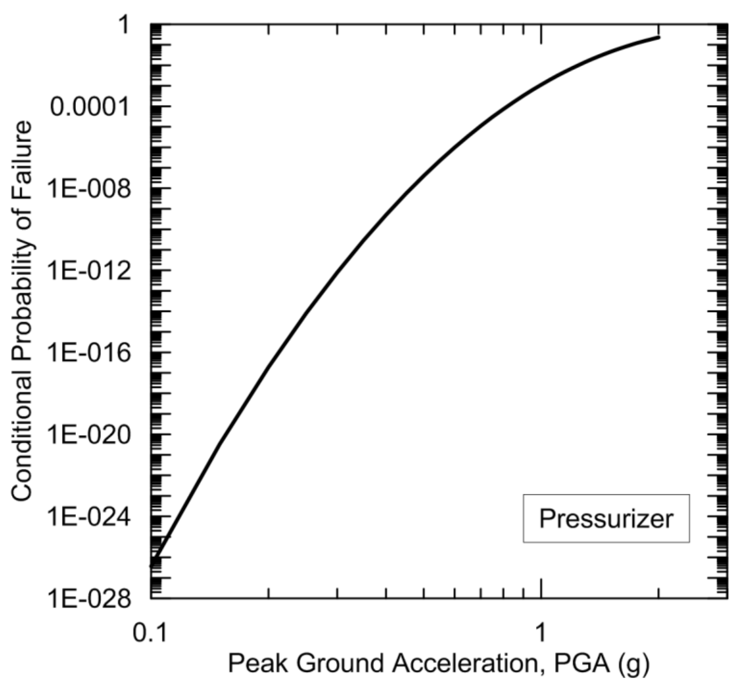

Fig. 7. Sample fragility curve for a pressurizer.

III. FAILURE PROBABILITY

The hazard curve is multiplied by the fragility curve to produce the failure probability curve, which is shown in Fig. 8. The figure shows that a change from a normal distribution to a student's t-distribution results in a lowering of the failure probability of the pressurizer for the PGA range of interest. When the area under the curves are calculated, the original GMPE results in a failure frequency of approximately 3.74×10^{-7} / year while the utilization of a t-distribution results in a failure frequency of approximately 3.02×10^{-9} / year. According to this, the t-distribution has reduced the probability of failure from an earthquake by at least 2 fold. Inclusion of PGAs below 0.1 g and above 2 g would most likely not decrease the gap in failure frequencies.

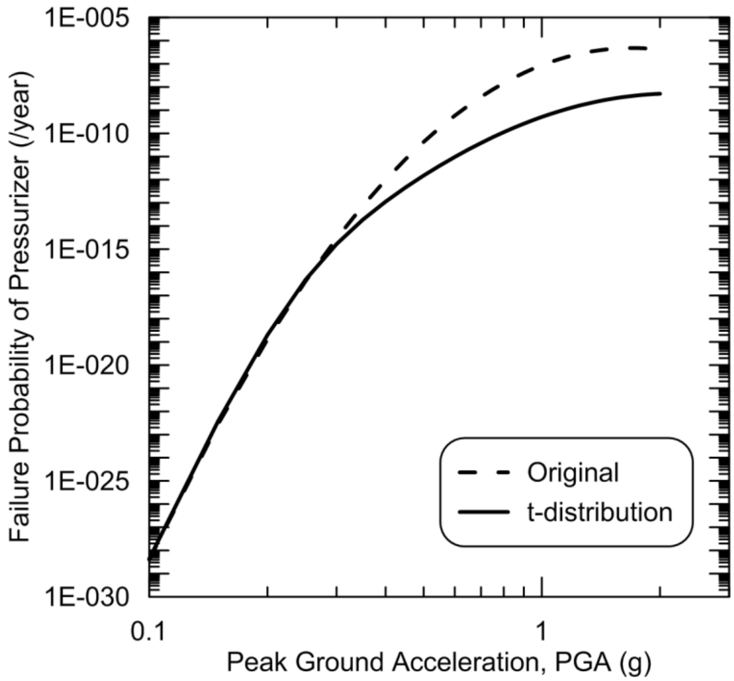

Fig. 8. Failure probability of a sample pressurizer.

IV. CONCLUSIONS

The use of a normal distribution to describe the randomness in the natural logarithm of earthquake PGA values in a GMPE is perhaps sometimes unwarranted given the data available and the environmental conditions of the data. This preliminary study attempted to address this issue by switching from a normal distribution to a student's t-distribution. A popular GMPE was selected to analyze a scenario of two faults near a site. This resulted in a seismic hazard curve that showed significantly less probability of exceedance, attributed to the relatively low t-distribution standard deviations calculated. The seismic hazard curves were then multiplied by a generic fragility curve for a sample pressurizer resulting in a failure probability curves. The failure probability curve based on a t-distribution plotted below the failure probability curve calculated from the original GMPE. Calculating the areas under the failure probability curves resulted in failure frequencies of 3.74×10^{-7} / year for the normal distribution and 3.02×10^{-9} / year for the adjusted t-distribution. Although these calculations were based on a PGA range of 0.1 to 2 g, it is suspected that expanding on this range would not make a significant difference in the magnitude of the difference between the two failure frequencies.

ACKNOWLEDGMENTS

This work was supported by the 2016 Research Funds of the KEPCO International Nuclear Graduate School (KINGS), Republic of Korea.

REFERENCES

1. I. M. IDRISS, "An NGA Empirical Model for Estimating the Horizontal Spectral Values Generated by Shallow Crustal Earthquakes," *Earthquake Spectra*, **24**, 1 (2008).
2. M. POWER, B. CHIOU, N. ABRAHAMSON, Y. BOZORGNIA, T. SHANTZ and C. ROBLEE, "An Overview of the NGA Project," *Earthquake Spectra*, **24**, 1 (2008).
3. B. CHIOU, R. DARRAGH, N. GREGOR, and W. SILVA, "NGA Project Strong-Motion Database," *Earthquake Spectra*, **24**, 1 (2008).
4. *Seismic Fragility Application Guide*, EPRI, Palo Alto, CA (2002).
5. *A Methodology for Analyzing Precursors to Earthquake-Initiated and Fire-Initiated Accident Sequences*, U.S. Nuclear Regulatory Commission, Report NUREG/CR-6544 (1998).