Kim, Yong Cheol; Kim, Byeong-Jun; Park, Jin Won; Cho, Yong-Sik

Tsunami Flooding Probability determined by Probability Distribution Type

TSUNAMI FLOODING PROBABILITY DETERMINED BY PROBABILITY DISTRIBUTION TYPE

Yong Cheol Kim[1], Byeong-Jun Kim[2], Jin Won Park[3] and Yong-Sik Cho[4]

ABSTRACT

Recently, catastrophic tsunamis triggered by impulsive undersea earthquakes have occurred frequently at subduction zones around the Pacific Ocean. These tsunamis have resulted in the loss of a huge number of lives and caused severe property damage. Thus, proper countermeasures to mitigate the tsunami damage should be made. For this reason, a tsunami hazard area at the Jumunjin Port in Korea is predicted by using a stochastic technique in this study. To find the hazard area, a numerical simulation of a virtual tsunami is conducted, and a goodness-of-fit test for the simulated flooding data is carried out. A best-fit probability distribution type for the flooding data is determined by a Probability Plot Correlation Coefficient (PPCC) test among the goodness-of-fit tests. The probabilities of flooding exceeding the criterion height are calculated by using a cumulative distribution function corresponding to the best-fit probability distribution type. The tsunami hazard area is estimated by these probabilities.

1. INTRODUCTION

In recent years, tsunamis triggered by undersea earthquakes have devastatingly attacked several coastal communities. The East Japan Tsunami Event occurred on March 11, 2011 and caused about 16,000 human deaths and 210 billion US dollars in property damage according to a report by NOAA (National Oceanic and Atmospheric Administration, 2012). These tsunamis are difficult to predict, and so it is difficult to prevent resultant loss of human lives and property damage. Thus, a proper countermeasure to mitigate the damage should be made.

The tsunami is difficult to reproduce by hydraulic modeling, and empirical studies on behaviors of tsunamis are very limited. Therefore, the studies on propagation and associated run-up process of tsunamis employ a numerical model to predict the flood area and height. In the present study, a numerical simulation of a tsunami is conducted with a numerical model developed by Cho *et al*. (2007). The numerical model uses the two-dimensional linear Boussinesq and nonlinear shallow-water equations as the governing equations for tsunami propagation and inundation,

[1] Graduate Student, Department of Civil and Environmental Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea (atdoyou@hanyang.ac.kr)
[2] Graduate Student, Department of Civil and Environmental Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea (astep4ch@hanyang.ac.kr)
[3] Graduate Student, Department of Civil and Environmental Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea (hanriver@paran.com)
[4] Corresponding Author, Professor, Department of Civil and Environmental Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea (ysc59@hanyang.ac.kr)

respectively. A group of virtual tsunamis is reproduced to obtain the flooding data. The earthquake fault parameters suggested by the KEDO (Korean Peninsula Energy Development Organization, 1999) are used.

To determine the best-fit probability distribution type for the tsunami flooding data, the goodness-of-fit test including the chi-square, Cramer-von Mises (CVM), Kolmogorov-Smirnov (K-S), and Probability Plot Correlation Coefficient (PPCC) test is conducted. The PPCC test has a more powerful rejection than the other tests, and can be applied easily among the goodness-of-fit tests. The test is used to find the best-fit probability distribution type in this study. The test was developed by Filliben (1975) to examine the normality of given data, and it has been applied to several probability distribution types. Recently, Choi and Jacobs (2007) calculated the PPCC test statistics of the normal and lognormal distribution to identify optimal soil moisture distributions in experiment fields. For several probability distributions, Heo *et al*. (2008) developed the regression equations calculating the PPCC test statistics from the number of data samples and significance level. In this study, the best-fit probability distribution type for the tsunami flooding data is determined by the PPCC test, and the probabilities of a flood exceeding a specific height is estimated to find a tsunami hazard area.

2. NUMERICAL MODEL

Because the free surface displacement of tsunamis is relatively less than the water depth in an ocean, the nonlinearity terms can be neglected for the simulation of tsunami propagation. The wave length of the tsunamis compared to the tide is relatively short and the tsunamis propagate over a long distance, however, the frequency dispersion effects may play important roles in tsunami propagation and should be considered in the governing equations. Thus, the linear Boussinesq equations are adequate as the governing equations of tsunami propagation (Cho *et al*., 2007). The linear Boussinesq equations are written in the eqs. 1-3.

$$\frac{\partial \zeta}{\partial t}+\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}=0 \tag{1}$$

$$\frac{\partial P}{\partial t}+gh\frac{\partial \zeta}{\partial x}=\frac{h^2}{2}\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}\left(\frac{\partial P}{\partial t}\right)+\frac{\partial}{\partial y}\left(\frac{\partial Q}{\partial t}\right)\right]-\frac{h^3}{6}\frac{\partial}{\partial x}\left[\frac{\partial^2}{\partial t\partial x}\left(\frac{P}{h}\right)+\frac{\partial^2}{\partial t\partial y}\left(\frac{Q}{h}\right)\right] \tag{2}$$

$$\frac{\partial Q}{\partial t}+gh\frac{\partial \zeta}{\partial y}=\frac{h^2}{2}\frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}\left(\frac{\partial P}{\partial t}\right)+\frac{\partial}{\partial y}\left(\frac{\partial Q}{\partial t}\right)\right]-\frac{h^3}{6}\frac{\partial}{\partial y}\left[\frac{\partial^2}{\partial t\partial x}\left(\frac{P}{h}\right)+\frac{\partial^2}{\partial t\partial y}\left(\frac{Q}{h}\right)\right] \tag{3}$$

where ζ is the free surface displacement, h is the still water depth, P and Q are the depth-averaged volume fluxes in the $x-$axis and $y-$axis directions, respectively, and g is the gravity acceleration.

Because the wave length of tsunami becomes shorter and the amplitude becomes larger as the tsunami propagates to a coastal region, the linear Boussinesq equations are no longer valid as the governing equations. In addition, the bottom friction effects become more important in the coastal region. Therefore, the nonlinear convective inertia force and bottom friction terms become increasingly important, while the frequency dispersion terms become less relevant, so the nonlinear shallow-water equations including bottom friction term should be used as the governing equations to simulate the tsunami inundation. The nonlinear shallow-water equations are represented in eqs. 4-6.

$$\frac{\partial \zeta}{\partial t}+\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}=0 \tag{4}$$

$$\frac{\partial P}{\partial t}+\frac{\partial}{\partial x}\left(\frac{P^2}{H}\right)+\frac{\partial}{\partial y}\left(\frac{PQ}{H}\right)+gH\frac{\partial \zeta}{\partial x}+\tau_x H=0 \tag{5}$$

$$\frac{\partial Q}{\partial t}+\frac{\partial}{\partial x}\left(\frac{PQ}{H}\right)+\frac{\partial}{\partial y}\left(\frac{Q^2}{H}\right)+gH\frac{\partial \zeta}{\partial y}+\tau_y H=0 \tag{6}$$

where H is the total water depth defined as $H=h+\zeta$, and τ_x and τ_y are the bottom friction terms represented by the Manning equation. The Manning equations can be written in the following form, where n is the Manning coefficient.

$$\tau_x=\frac{gn^2}{H^{10/3}}P\left(P^2+Q^2\right)^{1/2} \text{ and } \tau_y=\frac{gn^2}{H^{10/3}}Q\left(P^2+Q^2\right)^{1/2} \tag{7}$$

Eq. 4-6 are discretized by finite difference methods. A detailed description of the numerical scheme corresponding to eq. 4-6 can be found in many researches and is not repeated here (*Cho et al*., 2007; *Sohn et al*., 2009).

3. NUMERICAL SIMULATION

The Korean Peninsula Energy Development Organization (1999) identified 11 points that have a high probability of earthquake occurrence, one of which is used in this study. This location information and fault parameters are given in Table 1.

Table 1 Location information and fault parameters.

Location		***H***	***θ***	***δ***	***λ***	***L***	***W***	***D***	***M***
Long. (°*E*)	**Lat. (°*N*)**	**(*km*)**	**(°)**	**(°)**	**(°)**	**(*km*)**	**(*km*)**	**(*km*)**	**(Magnitude)**
139.0	41.7	1.0	1.0	40	90	125.899	62.945	6.31	8.0

In Table 1, H is the depth of fault plane, θ is the strike angle, δ is the dip angle, λ is the slip angle, L is the length of fault, W is the width of fault, D is the dislocation of fault, and M is the magnitude of earthquake.

Table 2 and Figure 1 show the computational information and domains for each region, respectively.

Table 2 Computational information for each region.

Region	Grid size ($\Delta x = \Delta y$, m)	Mesh number		Time step size (Δt, sec)	Type of numerical model
		x	y		
A	1215.0	970	1027	3.00000	Linear
B	405.0	1090	1228	1.00000	Linear
C	135.0	601	817	0.33333	Linear
D	45.0	604	652	0.11111	Linear
E	15.0	556	556	0.03704	Nonlinear
F	5.0	460	427	0.03704	Nonlinear

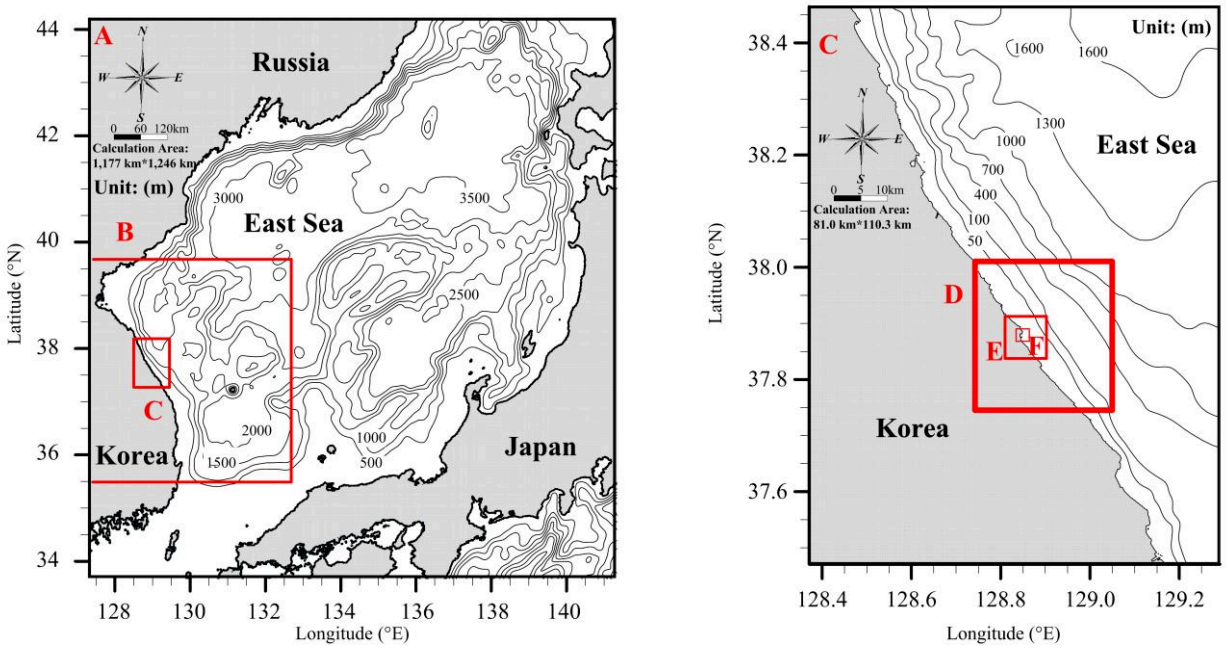

Figure 1 Computational domains for each region.

Table 3 shows the boundary conditions for each region. Region A used a free transmission condition, and the other regions used a dynamic linking method as a boundary condition for open sea. From region A to D fully reflected condition was used, and E and F regions used a moving boundary condition as a boundary condition for land.

Table 3 Boundary condition for each region.

Region	A	B	C	D	E	F
Boundary condition for land	Fully reflected	Fully reflected	Fully reflected	Fully reflected	Moving boundary	Moving boundary
Boundary condition for open sea	Free transmission	Dynamic linking	Dynamic linking	Dynamic linking	Dynamic linking	Dynamic linking

4. TSUNAMI FLOODING PROBABILITY

The PPCC test was applied to determine the probability distribution of the tsunami flooding data in this study, and the normal, log-normal, exponential and Gumbel distribution were assumed as the probability distributions. The test uses a correlation coefficient r_c for the ranked samples X_i and the fitted quantiles M_i. The correlation coefficient is calculated by eq. 8 (Filliben, 1975).

$$r_c = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(M_i - \overline{M}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \sum_{i=1}^{n}\left(M_i - \overline{M}\right)^2}} \tag{8}$$

where $\overline{X}$ and $\overline{M}$ are the mean values of the samples X_i and the fitted quantiles M_i, respectively, and n is the number of samples. The correlation coefficient is a measure of the linearity of probability plot. If the samples arise from an assumed distribution, the ranked samples versus the expected values in the assumed distribution will be approximately linear. Thus, if the correlation coefficient has a value near one, the observations will correspond to the assumed distribution. Therefore, the maximum correlation coefficient was chosen from among the correlation coefficients for each probability distribution to determine the probability distribution type instead of comparing the correlation coefficient with the PPCC test statistics. In other words, the probability distribution type having the maximum correlation coefficient was determined to be the probability distribution of a certain point.

To calculate the correlation coefficient, the fitted quantile should be obtained. The fitted quantile is calculated by using eq. 9.

$$M_i = \Phi^{-1}\left(P_i\right) \tag{9}$$

where Φ^{-1} is an inverse function of cumulative probability distribution, and P_i is a plotting position. The inverse functions of cumulative probability distribution for the normal, log-normal, exponential and Gumbel distributions were used in this study. The inverse functions of each cumulative probability distribution are shown in Table 4.

Table 4 Inverse functions of cumulative probability distribution for each distribution.

Probability distribution	**Inverse function of cumulative probability distribution**
Normal	$z_x = 5.0633\left[F^{0.135} - \left(1-F\right)^{0.135} \right]$
Log-normal	$z_y = 5.0633\left[F^{0.135} - \left(1-F\right)^{0.135} \right]$
Exponential	$x = -\frac{1}{\eta}\ln\left(1-F\right)$
Gumbel	$x = \xi - \alpha \ln\left[-\ln\left(F\right) \right]$

In Table 4, x is a random variable, z_x and z_y are the normalized random variables for the normal and log-normal distribution, F is a quantile, and η, ξ, α are the shape parameters, the location parameter, the scale parameter, consecutively.

Because the plotting position plays a significant role in estimation of the correlation coefficient, the proper plotting position should be considered. The plotting position formulas of Blom (1958) and Gringorten (1963) were used to calculate the proper fitted quantiles in this study, and represented in eqs. 10-11, respectively.

$$P_i = \frac{i-0.375}{n+0.25} \tag{10}$$

$$P_i = \frac{i-0.44}{n+0.12} \tag{11}$$

where i is a rank. The plotting positions for the normal, log-normal, exponential distribution were calculated by using an eq. 10, and the other was calculated by using an eq. 11.

After the probability distribution was determined, the probability that flooding will exceed the criterion height was calculated by using the cumulative distribution function corresponding to each probability distribution. The cumulative distribution functions used in this study are listed in Table 5. In Table 5, $F(x)$ and $F(y)$ are the cumulative probability, μ_x and μ_y are the mean of x and y, σ_x and σ_y are the standard deviation of x and y, respectively. The criterion height was selected as 30 cm in this study. Finally, the probability that flooding will exceed the criterion height was obtained at each point.

Table 5 Cumulative distribution function.

Probability distribution	Cumulative distribution function
Normal	$F(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{1}{2\sigma_x^2}(x-\mu_x)^2}dx$
Log-normal	$F(y)=\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}\sigma_y}e^{-\frac{1}{2\sigma_y^2}(y-\mu_y)^2}dy$
Exponential	$F(x)=1-\exp(-\eta x)$
Gumbel	$F(x)=\exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right]$

5. RESULTS AND DISCUSSION

Figure 2 represents the maximum flooding heights obtained by numerical simulation for tsunami. The values of Figure 2 mean the maximum flooding height during the numerical simulation period. The northern area and near the coastline generally have a higher flood height than other flooded areas. As shown in Figure 2, the maximum flooding height is approximately 1.2 *m*.

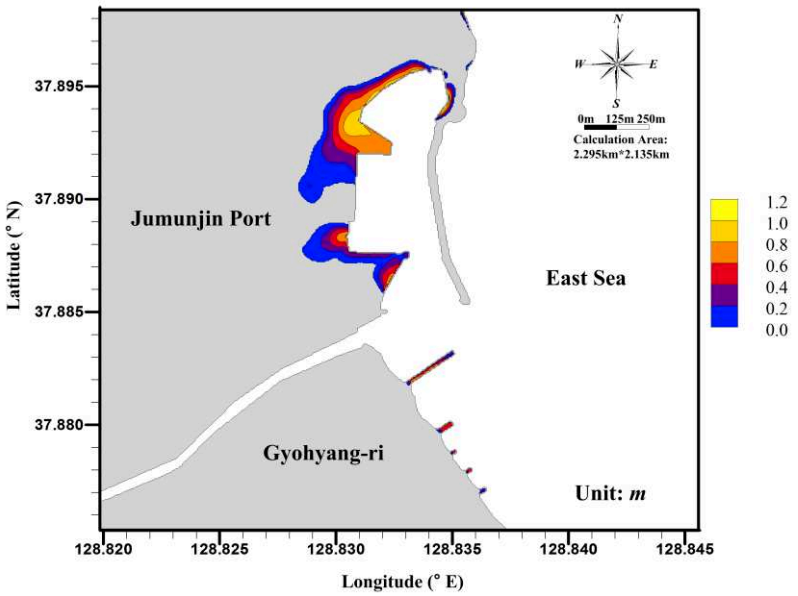

Figure 2 The maximum flooding heights.

A goodness-of-fit test, namely the PPCC test, was conducted at all points where flooding occurred, and the correlation coefficients for the 4 types of probability distribution (Normal, Log-normal, Exponential, Gumbel distribution) were calculated at each point. The largest correlation coefficient value calculated at each point is shown in Figure 3, which shows that correlation coefficients higher than 0.95 were obtained at most of the points where flooding occurred. That the

value of maximum correlation coefficient is approximately 1.0 indicates that the probability distribution type agrees with at least one of the 4 assumed probability distribution types. Thus, the result of Figure 3 indicates the usefulness of maximum correlation coefficient concept. The probability distribution type with the highest correlation coefficient at each point was determined to be the probability distribution type for that point.

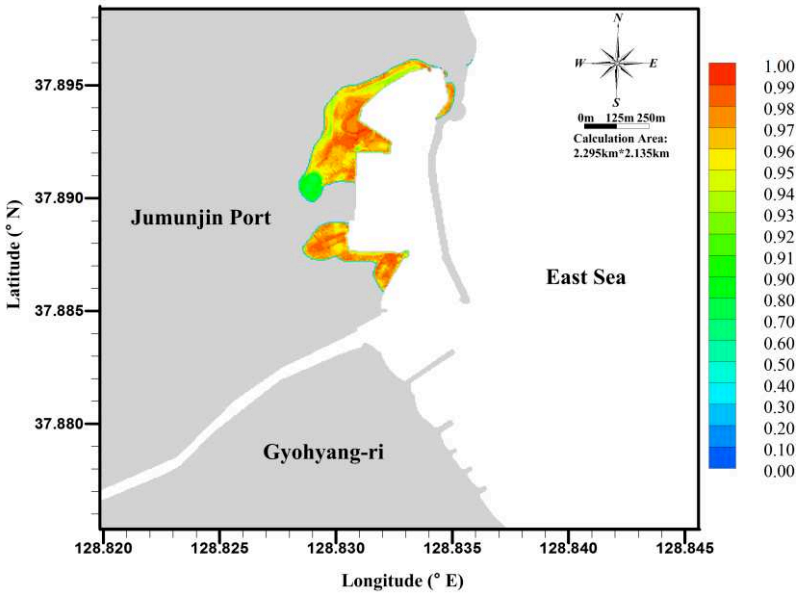

Figure 3 The maximum correlation coefficients.

The probability of a flood exceeding 30 *cm* was calculated by using the cumulative distribution function for the selected probability distribution type at each point, and the result is shown in Figure 4. In Figure 4, the northern area has a flooding probability greater than 0.6, and these points correspond with the lowest-elevation areas of Jumumjin Port. These results showed good agreement with the maximum flooding heights shown in Figure 2. When comparing the maximum flooding heights of Figure 2 with the model results, the points shown to have a high flood height generally correspond to those with a high flooding probability. The result of Figure 4 indicates the degree of risk for the virtual tsunami inundation.

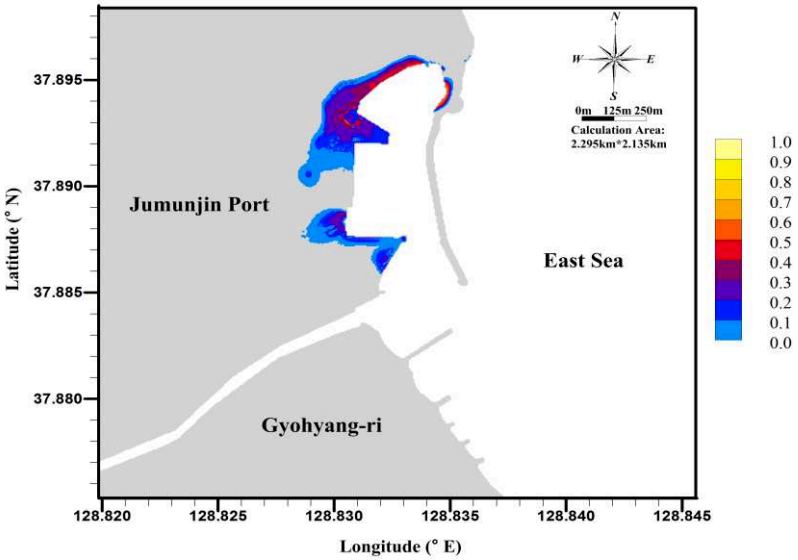

Figure 4 The flooding probabilities in Jumumjin Port.

ACKNOWLEDGEMENTS

This study was financially supported by the Construction Technology Innovation Program(08-Tech-Inovation-F01) through the Research Center of Flood Defense Technology for Next Generation in Korea Institute of Construction & Transportation Technology Evaluation and Planning(KICTEP) of Ministry of Land, Transport and Maritime Affairs(MLTM)

REFERENCES

Blom, G. (1958) Statistical Estimates and Transformed Beta-Variables, Wiley, New York.

Cho, Y.-S., Sohn, D.-H. and Lee, S.-O. (2007) "Practical Modified Scheme of Linear Shallow-Water Equations for Distant Propagation of Tsunamis", Ocean Engineering, Vol. 34, No. 11-12, pp. 1769-1777.

Choi, M. and Jacobs, J.M. (2007) "Soil Moisture Variability of Root Zone Profiles within SMEX02 Remote Sensing of Footprints", Advances in Water Resources, Vol. 30, No. 4, pp. 883-896.

Filliben, J.J. (1975) "The Probability Plot Correlation Coefficient Test for Normality", Technometrics, Vol. 17, No. 1, pp. 111-117.

Gringorten, I.I. (1963) "A Plotting Rule for Extreme Probability Paper", Journal of Geophysical Research, Vol. 68, No. 3, pp. 813-814.

Heo, J.-H., Kho, Y.W., Shin, H., Kim, S. and Kim, T. (2008) "Regression Equations of Probability Plot Correlation Coefficient Test Statistics from Several Probability Distributions", Journal of Hydrology, Vol. 355, No. 1-4, pp. 1-15.

Korean Peninsula Energy Development Organization. (1999) Estimation of Tsunami Height for KEDO LWR Project, Korea Power Engineering Company, Inc., Korea.

National Oceanic and Atmospheric Administration. (2012) Tsunami Event Database, http://www.ngdc.noaa.gov/hazard/hazards.shtml (accessed July 24, 2012).

Sohn, D.-H., Ha, T. and Cho, Y.-S. (2009) "Distant Tsunami Simulation with Corrected Dispersion Effects", Coastal Engineering Journal, Vol. 51, No. 2, pp. 123-141.