11th International Workshop on Ship and Marine Hydrodynamics
Hamburg, Germany, September 22-25, 2019

Prediction on Extreme Distribution of Sloshing Loads Considering Various Statistical Models and Threshold Values

Jeoungkyu[1] Lee, Jieung Kim[1], Yangjun Ahn[1], and Yonghwan Kim[1]*

[1]Seoul National University, Department of Naval Architecture and Ocean Engineering
1, Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea

*Corresponding author, yhwankim@snu.ac.kr

ABSTRACT

In this study, the peak-over-threshold method is applied to investigate the influence of the threshold used in the statistical analysis of sloshing loads. Weibull distribution, generalized Pareto distribution, and log-logistic distribution are used to the statistical analysis. In addition to the previously used moment method, the parameters of distributions are estimated using the L-moment method and the maximum likelihood function estimation. Statistical analysis is performed based on the experimental data obtained from 20 repeated experiments of 5 hours in real scale. In order to compare the results between statistical models, the probability plot correlation coefficient and the 3-hour most probable maximum are used. The results of this study can be one of the indicators for the statistical analysis for sloshing load prediction according to experimental conditions.

1 INTRODUCTION

Sloshing is a phenomenon in which the internal fluid fluctuates due to the external force applied to the cargo hold. The sloshing impacts to the cargo hold, which may cause damage to the structure. A study on the impact load estimation of such sloshing is becoming more active as the demand of LNG carrier is increasing. Numerical approaches to sloshing phenomena are increasing due to advances in computer performance and computational methods. However, due to the strong nonlinearity of the sloshing flow, experimental studies are now more active and more reliable. These experimental results are widely used for sloshing load estimation and numerical simulation.

The impact pressure values obtained from the sloshing model test are used to estimate the design sloshing load through statistical processing. Mathiesen[1] and Gran[2] conducted a basic study on the statistical approach for sloshing load estimation. Mathiesen applied the Weibull distribution to analyse the maximum impact pressure obtained from arbitrary pitch motions, and Gran compared the results of applying the Weibull distribution and Frechet distribution. In addition, Grazcyz et al.[3] used the Weibull distribution and the Generalized Pareto distribution for statistical analysis of sloshing model experiments for 5 hours in real scale. After that, Kuo et al.[4] conducted a study on maximum pressure estimation and confidence interval estimation through statistical model setting for LNG sloshing problem. Kim et al.[5] analysed the effect of extraction and smoothing method on experimental data by applying the Weibull distribution and the generalized Pareto distribution. Filon et al.[6] applied the 3-parameter Weibull distribution, the generalized Pareto distribution, and the generalized extreme value distribution to the experimental results, and compared the results with the Kolmogorov Smirnov evaluation method. Recently, Cetin et al.[7] applied various models such as the Weibull distribution, the generalized Pareto, generalized extreme value, and log-logistic distribution to the results of sloshing experiments, which suggested the applicability of other statistical models.

Predicting the accurate maximum peak pressure in a designated return period is a very important factor in the structural design stage of the LNG cargo hold. At this time, the estimated maximum impact pressure values are highly dependent on the statistical model applied to the experimental results. In the short-term prediction, the difference of the estimated values according to the statistical model is not large, but in the long-term prediction, the difference of the estimated values increases according to the selection of the distribution function. According to the guidelines presented by the Societies, the Weibull distribution and the generalized Pareto applying the moment method as the statistical analysis model are presented the most, but there is no clear basis for why such distributions are used. In this study, a variety of distribution functions and parametric estimation methods to the analysis of sloshing test results are applied to find out whether a more general or accurate statistical model can be defined in estimating the sloshing impact pressure value.

In the real ship design, high sloshing impact pressure values are considered to be important, but due to the high impact rate of the low impact pressure values, there is a possibility that the estimated impact pressure value is deviated to the low impact pressure. Therefore, in this study, the peak-over-threshold (POT) method usually used in the generalized Pareto distribution is applied to other statistical models to compare the extreme loads according to the threshold value.

Statistical analysis is performed based on the experimental data obtained by repeating the experiment corresponding to the 5 hours in real scale 20 times. In order to compare the results between the statistical models applied in this study, the probability plot correlation coefficient (PPCC) and the 3-hour most maximum probable value are used.

2 STATISTICAL MODEL

2.1 Distribution functions

Weibull distribution, generalized Pareto distribution, and log-logistic distribution are applied to estimate maximum impact pressure. The cumulative probability distribution of each statistical model is as follows:

Three-parameter Weibull distribution (WBL)

$$F\left(x|\gamma,\alpha,\beta\right)=1-\exp\left(-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right) \tag{1}$$

Generalized Pareto distribution (GP)

$$F\left(x|\gamma,\beta\right)=1-\left(1+\frac{\gamma x}{\beta}\right)^{-\frac{1}{\gamma}} \tag{2}$$

Three-parameter log-logistic distribution (LL)

$$F\left(x|\gamma,\alpha,\beta\right)=\left(1+\left(\frac{\beta}{x-\alpha}\right)^{\gamma}\right)^{-1} \tag{3}$$

In the distribution functions, γ is the shape parameter, α is the location parameter, and β is the scale parameter. For the variable x, it should be greater than or equal to the location parameter. In case of the generalized Pareto distribution, there is no location parameter, and the peak-over-threshold (POT) method is applied to improve the accuracy of the approximation.

The POT method is a method of performing distribution function approximation based on data over the threshold value. In general, the distribution approximation is performed considering the peak values corresponding to the upper 8% of data. However, in this study, the approximation results according to the respective threshold values are compared by dividing the upper 5% to the 50% of the data into 5% units.

2.2 Distribution fitting methods

It is important to apply appropriate methods to each distribution when estimating parameters. Because of the difference in parameters, the shape of the distribution function changes and a large difference in the maximum impact pressure estimation value may appear. In this study, to estimate the parameter values of each statistical model, the method-of-moments (MOM), the L-moment method (LMOM), the maximum likelihood estimation method (MLE) are applied.

The MOM is a method of estimating the parameters by correlating the mean, variance, and vorticity of statistical models with those of extracted data. The limitation of the MOM is that the range of shape parameters is limited in some statistical models such as generalized extreme value distribution.

As an alternative to the limit of the MOM, Hosking[8] proposed the LMOM in 1990. The LMOM is similar to the conventional moment method, but it is estimated by comparing the linear combination of the order statistics. This method estimates parameters by associating the L-moment and L-moment ratio of the approximated distribution with those of the extracted data.

The MLE is a method developed by R.A. Fisher in the 1920s and, is a parameter estimation method to find the distribution that makes the data most probable by maximizing the likelihood function.

Based on the study of statistical model analysis for the sloshing impact pressure by Cetin, the high approximated distributions with the method of parameter estimations are selected. In this study, several parameter estimation methods are applied to three distributions which are considered to have relatively high reliability. Table 1 summarized the methods and the brief notations used in this study.

Table 1 Parameter estimation method for distribution and notation

Distribution	Fitting method	Notation
Weibull distribution	Method-of-moments	WBL-MOM
	L-moments method	WBL-LMOM
Generalized Pareto distribution	Method-of-moment	GP-MOM
	Maximum likelihood estimation	GP-MLE
Log-logistic distribution	Method-of-moments	LL-MOM
	L-moments method	LL-LMOM

2.3 Goodness-of-fit

In this study, two methods are used to verify whether each statistical model closely approximates the experimental results. The first one is the probability plot correlation coefficient (PPCC) test. The PPCC test was first proposed by Filiben (1975) and later developed to be applicable to other distributions as well as normal distributions. In this method, a coefficient r indicating the relationship between the observed value X_i and the estimated displacement value M_i is defined, which is expressed by equation (4). Fundamentally, r provides a quantitative assessment of fitness by measuring the linearity of the probability plot (Heo et al.[9]). As r approaches 1, the experimental value approximate to the value obtained from the distribution.

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(M_i - \overline{M}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \sum_{i=1}^{n}\left(M_i - \overline{M}\right)^2}} \tag{4}$$

$$M_i = \Phi^{-1}\left(p_i\right) \tag{5}$$

$$p_i = 1 - \left(0.5\right)^{\frac{1}{n}}, i = 1 \tag{6}$$

$$p_i = \frac{i - 0.3175}{n + 0.365}, i = 2, \ldots, n-1 \tag{7}$$

$$p_i = \left(0.5\right)^{\frac{1}{n}}, i = n \tag{8}$$

$\bar{X}$ and $\bar{M}$ denote the mean values of the observations X_i and the fitted quantiles M_i, respectively and n is the sample size. In the equation of the estimate of order statistic median for M_i, Φ^{-1} is the inverse of cumulative distribution function by Cunnane[10].

The second method is a comparison of the results using the Maximum Probable Maximum (MPM) value. Since the PPCC test measures the linearity of the probability plot, it is necessary to complement the analysis of the sloshing impact pressure result with strong nonlinearity. Also, it is important to estimate the impact pressure at the desired probability because the analysis is focused on the high impact pressure when designing the cargo hold in real scale. The maximum estimated pressure value is obtained by calculating the maximum pressure corresponding to a certain probability in the excess probability distribution, and mainly uses the maximum estimated pressure value based on probability of 3 hours (hereinafter referred to as 3-hour MPM). In the case of the Weibull distribution, the probability α that the sloshing maximum impact pressure value exceeds p_α appears as follows.

$$\alpha = 1 - F(P_\alpha) = \exp\left(-\left(\frac{p_\alpha - \delta}{\beta}\right)^\gamma\right) \tag{9}$$

p_α can be summarized as follows

$$p_\alpha = \delta + \beta\left[\log\left(\frac{1}{\alpha}\right)\right]^{\frac{1}{\gamma}} \tag{10}$$

Since extreme value estimates of the distribution function are usually computed on a time unit basis, the number of sample extractions N for m hours is:

$$N = \frac{3600nm}{t} \tag{11}$$

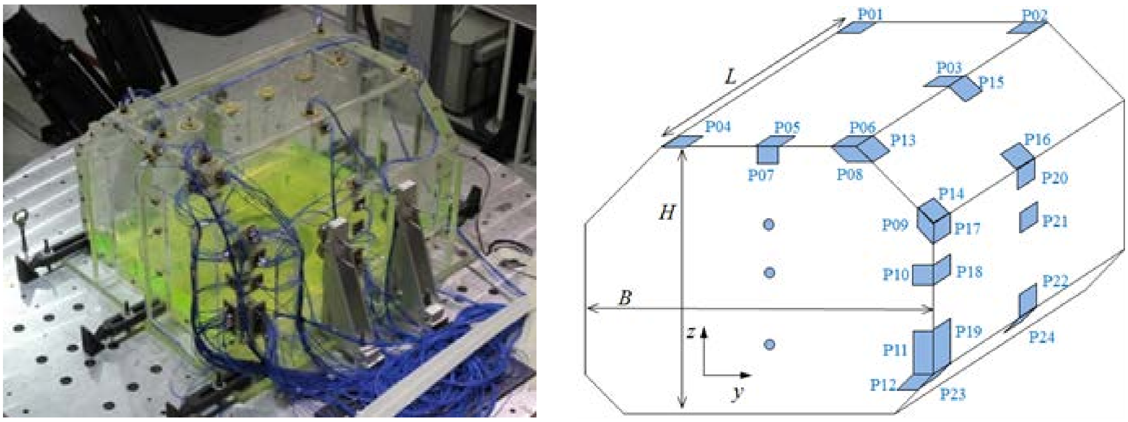
In equation (11), t is the total time of the data in real scale(sec), and n is the maximum number of times during the total time. The maximum estimated pressure obtained from this is calculated as follows.

$$p_{\max} = \delta + \beta\left[\log(N)\right]^{\frac{1}{\gamma}} \tag{12}$$

The maximum estimated pressure can be calculated in a similar way to the generalized Pareto distribution and the log-logistic distribution.

3 EXPERIMENTAL SETUP

The sloshing model experiment used in this study was carried out at Seoul National University and the sloshing model tank was installed on the 6 degree-of-freedom motion platform. The size of the model tank is 1/50 scaled down from real scale cargo hold, and the length (L) is 868.2 mm, the width (B) is 750 mm, and the height (H) is 555 mm. In order to avoid the effect of hydro-elasticity at the wall, the wall was made of 35-40 mm thick acrylic plate. Figure 1 shows a sloshing model cargo hold installed on the platform and a cluster panel layout on which the pressure sensor is placed.



(a) Experimental set up (b) Layout of sensor cluster panels

Figure 1. Sloshing model tank

20 times repeated experiments were carried out for the loading conditions of 20%, 50%, and 95% of the tank height. The significant wave height and mean wave period of the each sea condition for the repeated case were same, but the phases of wave were randomly generated. Table 2 summarizes the experimental conditions.

Table 2 Test conditions

Case	Filling depth	Heading angle	Sea state	Test time
Case 1	0.95*H*	150	Tz=9.5s Hs=12.5m	5hrs*20
Case 2	0.50*H*	90	Tz=11.5s Hs=9.5m	5hrs*20
Case 3	0.20*H*	90	Tz=7.5s Hs=7.5m	5hrs*20

In order to statistically analyze the sloshing load, impact pressure is identified using the peak-over-threshold (POT) method proposed by the international classification society. Figure 2 shows the method of extracting the peak pressure value. Among all the sensors in the same time zone, the largest peak pressure value is extracted within the moving time window range. The threshold value set in the experiment is 2.5 kPa, and the size of time window for the peak pressure value detection is 0.2 sec, which are empirical values. [5] The high-pass filter is used to remove unnecessary information of measured pressure signal for the analysis. After filtering the frequency components lower than 50 Hz and detecting the peak pressure values, peak values are sorted in descending order.

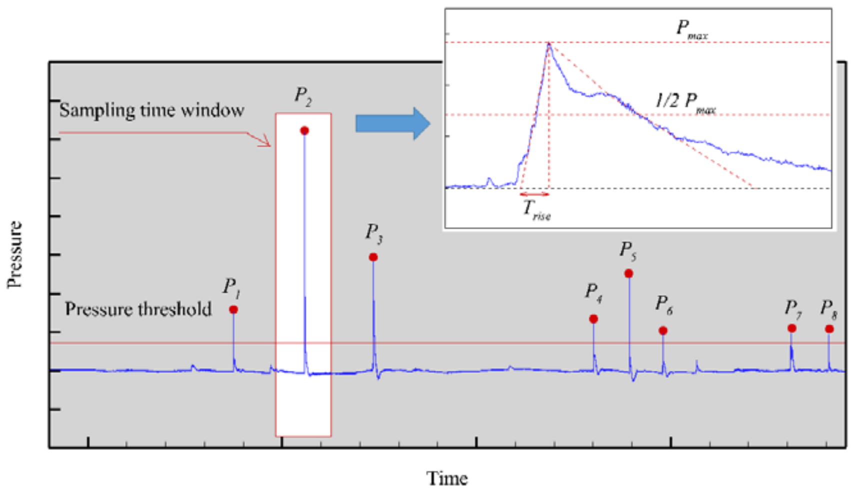


Figure 2. Peak-over-threshold (POT) method

4 RESULTS

4.1 Pressure signal

Figure 3 shows the part of pressure signal in time histories. The pressure in the *y*-axis is normalized by the density of the fluid (ρ), the acceleration of gravity (g), and the height of the cargo hold (H). Black dots are peak values that exceed the threshold and are values that are extracted for statistical analysis.

Figure 4 is the largest impact pressure signal for a representative case of the experiments under each loading condition. For ease of signal analysis, the point at which the maximum impact occurs is shifted to 0.01 second, and panel containing the channel with the highest impact pressure measured is labeled on the legend. Various types of pressure signals are observed according to the form of impact. In the case of high filling depth, oscillation due to air pocket generated at corners occurs, and under low filling depth, the impact pressure signal of the flip though type impact that rises from the wall surface is frequently observed.

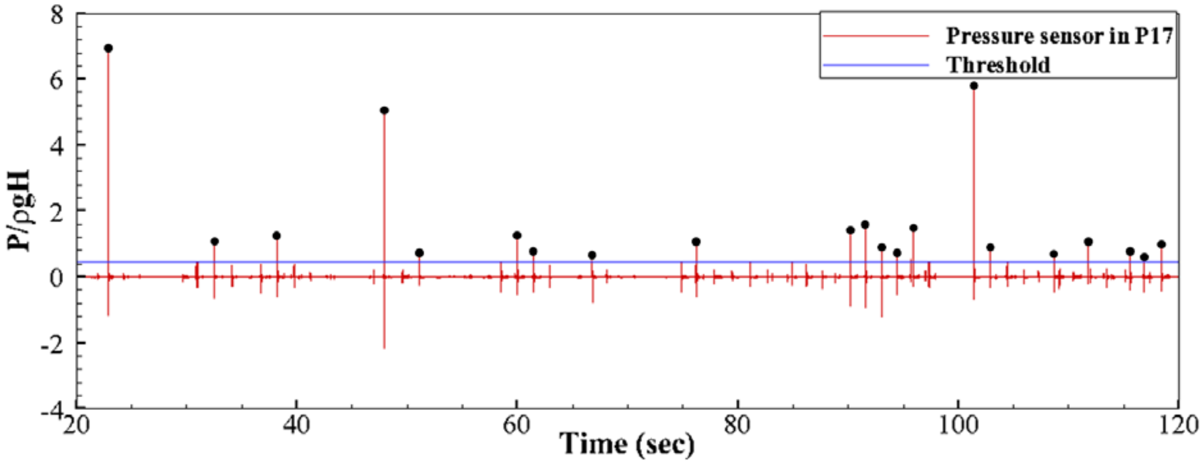
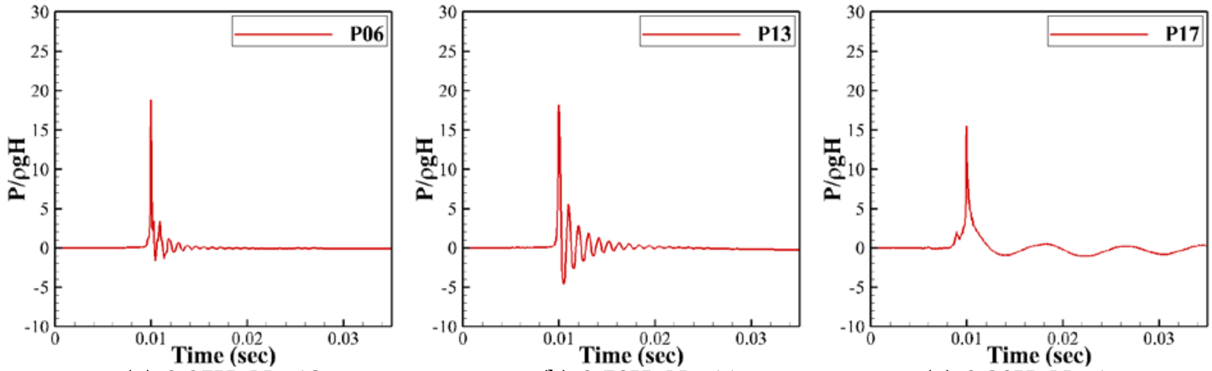


Figure 3. Pressure signal in time histories (0.20H, No.1)



(a) 0.95H, No.18 (b) 0.50H, No.11 (c) 0.20H, No.1

Figure 4. Single pressure peak in time histories

4.2 Plots of estimated distribution of statistical models

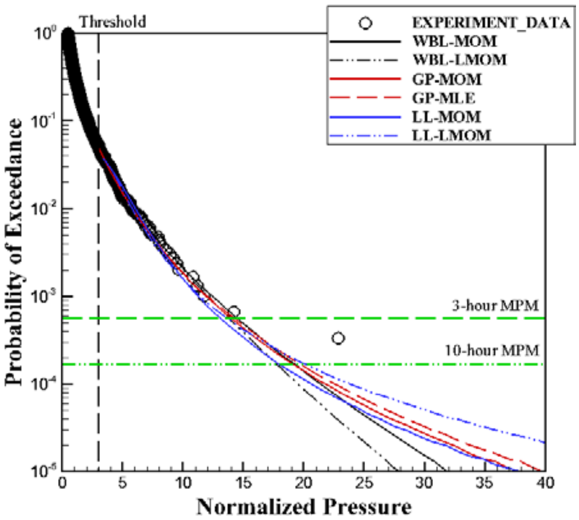
For each experimental result, POT method was applied by dividing the threshold value ratio from the upper 5% to 50% at intervals of 5%. Here, threshold value ratio means the ratio of data used for statistical analysis among whole data The cumulative probability distribution of the distribution approximation using the POT method is as follows.

$$F(x) = F_X(x)[1 - \hat{F}(u)] + \hat{F}(u) \quad (x > u) \tag{13}$$

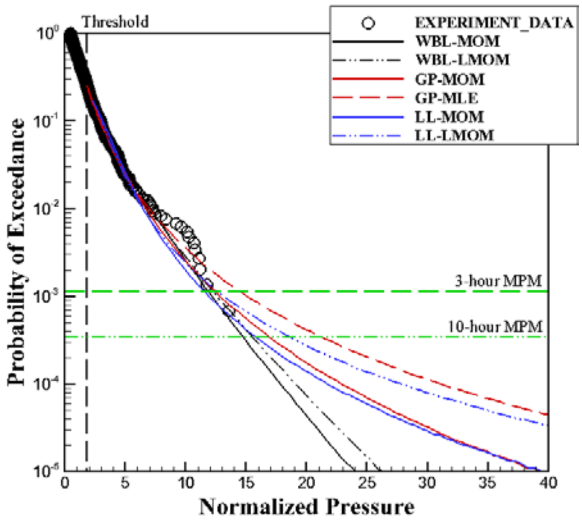
In equation (13), u represents the threshold value, and $F_X(x)$ is the cumulative probability distribution approximated by the data above the threshold value, and $\hat{F}(u)$ is an estimated value of $F(u)$, which is the number of data that is less than or equal to the threshold value in the entire sample divided by the total number of samples. (Pickands[11])

Figure 4 shows the cumulative probability distributions obtained by applying the POT method to each statistical model. Figure 5-(a) shows the result of the 9th repeated experiment under the 95% loading conditions, with a threshold value corresponding to the upper 5%, which is briefly indicated as 0.95H, No.9, and Thd 5%. The black dotted line in the longitudinal direction indicates the threshold value, the green dotted line in the horizontal direction indicates the 3-hour MPM, and the green dotted chain line indicates the 10-hour MPM. The pressure in the x-axis is a dimensionless value with the density of the fluid (ρ), the acceleration of gravity (g), and the height of the cargo hold (H).

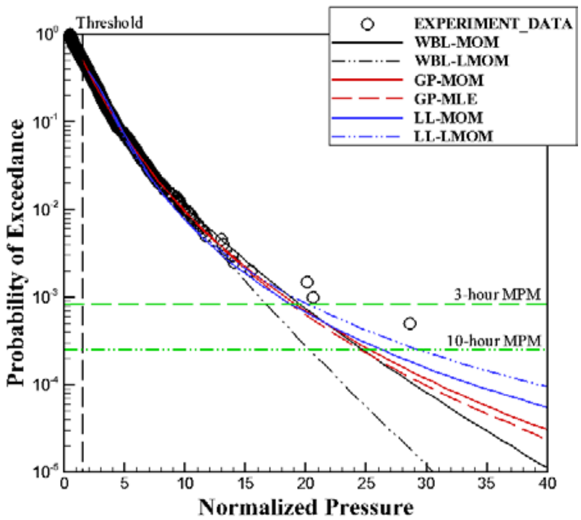
Figure 5-(a) shows that the approximate distribution of each statistical model is similar to the vicinity of 3-hour MPM when the cumulative probability distribution variation of the experimental results is not large. As shown in Figure 5-(b), when there is a variation in the cumulative probability distribution of the experimental results, there is a difference in the approximate distribution shape according to the statistical model. In the case of (c) in Figure 5, the cumulative probability distribution variation of the experimental results is small, but the impact pressure of low probability is large, and it can be seen that there is a difference in the approximate distribution shape between statistical models in the vicinity of 10-hour MPM.



(a) 0.95H, No.9, threshold ratio 5%



(b) 0.50H, No.12, threshold ratio 25%

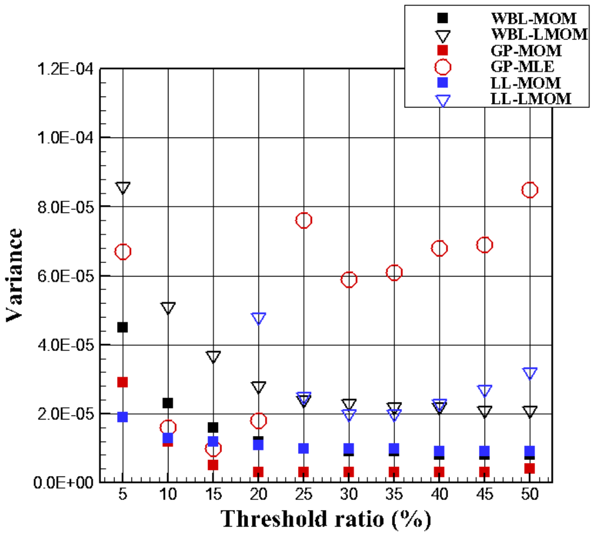


(c) 0.20H, No.18, threshold ratio 50%

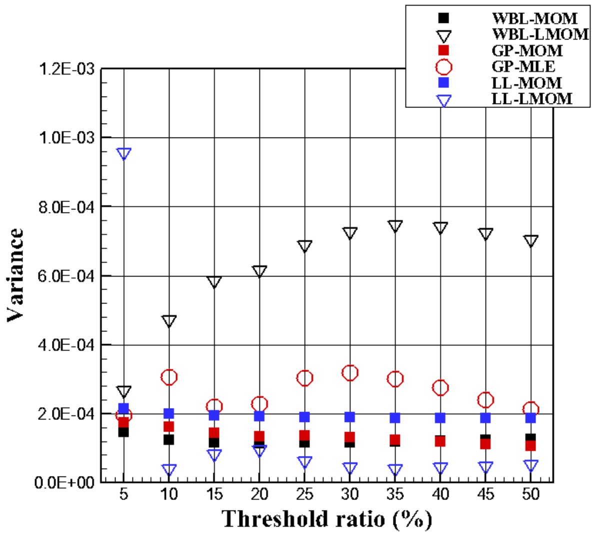
Figure 5. Probability of exceedance of representative conditions

4.3 Result of PPCC test

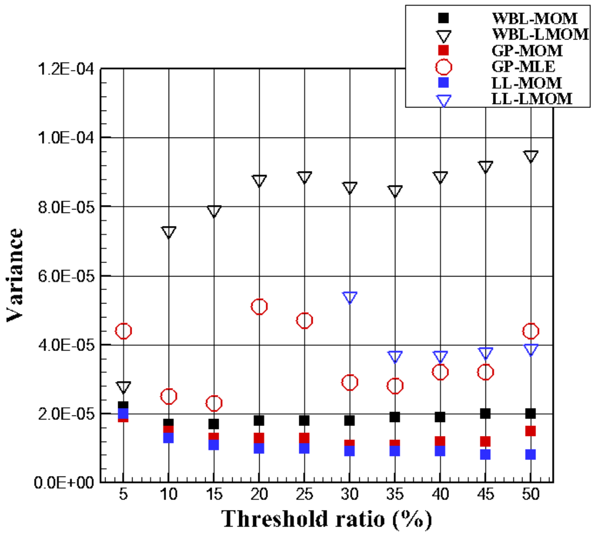
The dispersion of the results of the PPCC test for the experiments repeated 20 times, and the dispersion results according to the threshold value ratio are shown in Figure 6. In this case, the smaller the variance, the closer the PPCC test result is to 1.



(a) 0.95H



(b) 0.50H



(c) 0.20H

Figure 6. Variance of PPCC coefficient

In the loading condition of 0.95H, the dispersion tends to decrease as the threshold value increases. GP-MOM and WBL-MOM showed the least variance of PPCC test results. When the threshold ratio is 5-20%, the dispersibility of LL-LMOM is the highest. When the threshold ratio is 25-50%, GP-MLE is the most dispersed. On the other hand, the dispersion of GP-MOM and WBL-MOM is smaller than that of LL-LMOM at the loading condition of 0.50H. Finally, in the loading condition of 0.20H, the dispersion of LL-MOM is the smallest, followed by GP-MOM and WBL-MOM.

In the statistical model using the method-of-moment, it can be seen that the fluctuation according to the threshold value ratio is not when the threshold value ratio is larger than 15%. It is estimated that when the threshold value ratio is 5-10%, the number of data samples over the threshold value is small, and the variation according to the threshold value ratio is relatively large. It is presumed that when the threshold ratio is 5- 10%, the number of data samples exceeding the threshold value is small, and the fluctuation according to the threshold value ratio is relatively large. In the case of the L-moment method and the maximum likelihood estimation method, the approximation accuracy is lowered near the probability of high impact pressure, and the dispersibility is affected by the pressure distribution shape of the experimental result.

The statistical models with the least variance are different according to loading conditions. In general, the PPCC results of GP-MOM and WBL-MOM are close to 1, and the PPCC results of WBL-LMOM and GP-MLE are far from 1.

4.4 Comparison of most probable maximum

In the comparison of the maximum estimated pressure, the maximum value obtained by applying the 100-hour impact pressure data accumulated by the experiment repeated 20 times was set as a standard. Figure 7 shows the 3-hour MPM of the 100-hour cumulative impact pressure data. The remaining statistical models except for the generalized Pareto distribution were calculated using the conventional approximation method without the POT method.

As the filling level decreases, the overall 3-hour MPM increases. For the 0.95H and 0.20H of filling conditions, the 3-hour MPM of the statistical models except WBL-LMOM and LL-LMOM are similar to each other. At 0.50H filling condition, the 3-hour MPM of the statistical models except WBL-LMOM are similar.

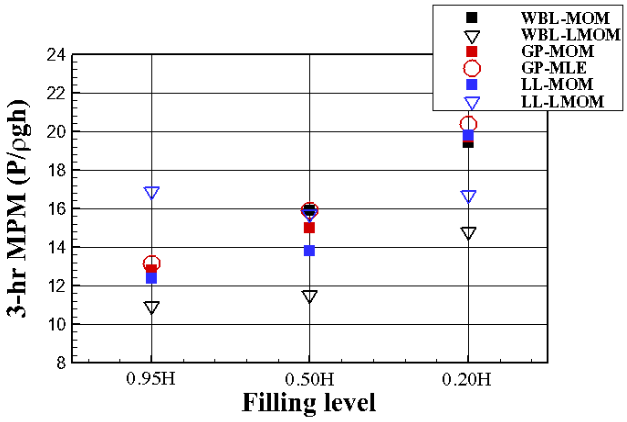


Figure 7. 3-hour MPM of 100-hours cumulative pressure data

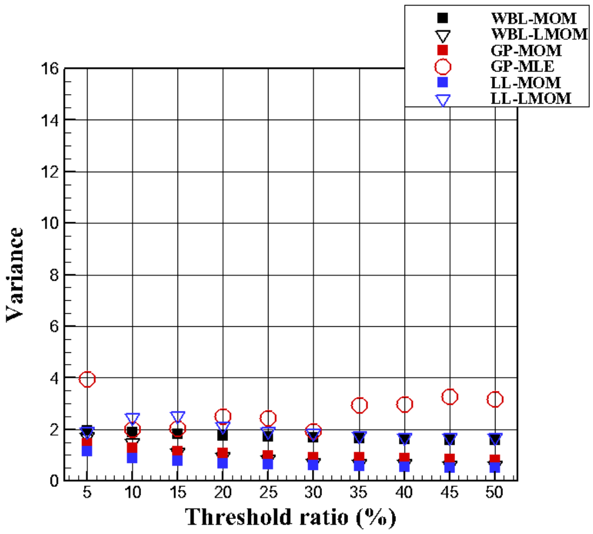
The distributions of the 3-hour MPM for the experiments repeated 20 times according to the threshold ratio are shown in Figure 8. In this case, the smaller the variance, the closer the 3-hour MPM results of the 5-hour experiment to the 3-hour MPM results of the 100-hour cumulative pressure data for each statistical model.

When the filling level is high, the range of free surface movement is not so large, so the impact accompanied by air pocket often occurs and the sloshing impact shape is similar. On the other hand, since the shape of the traveling wave mainly occurs at the lower filling level, the shape of the free surface becomes unstable, causing wave breaking or jump phenomenon, and various types of sloshing impacts appear. Therefore, as filling level decreases, the dispersion of peak pressure increases.

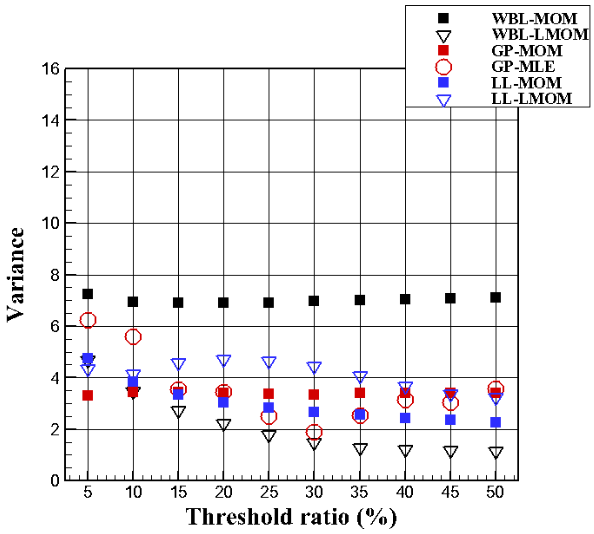
For the statistical model using the moment method, the variation range of dispersion according to the threshold ratio is smaller than that of the L-moment method or the maximum likelihood estimation method, and the dispersibility decreases as the threshold ratio increases. In particular, GP-MOM using the existing POT method also has a decrease in dispersibility as the threshold value ratio increases, which requires further examination of the upper 8% threshold value used in the past.

In the 0.95H filling level condition, low dispersibility is observed in the order of LL-MOM, WBL-LMOM and GP-MOM, and GP-MLE shows the highest dispersibility. In 0.50H filling level condition, low dispersibility is observed in the order of WBL-LMOM, LL-MOM, GP-MOM, and WBL-MOM shows the highest dispersibility. Finally, at the filling level of 0.20H, LL-MOM and WBL-MOM show the lowest dispersibility and dispersion of LL-LMOM is the highest.

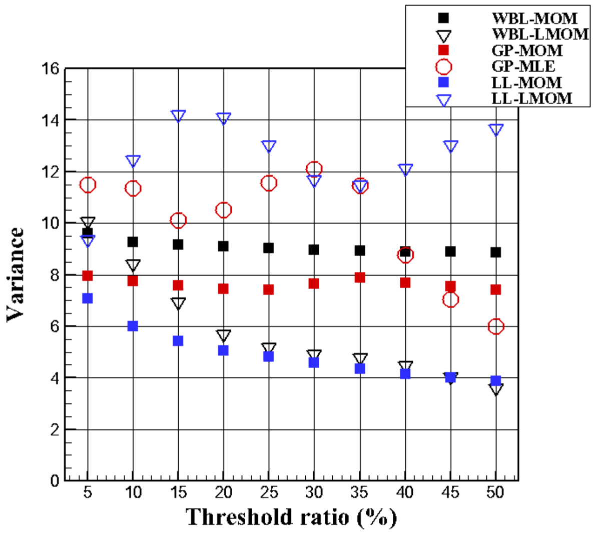
Taken together, the 3-hour MPM comparison shows that LL-MOM and WBL-LMOM have the lowest dispersibility. However, WBL-LMOM shows high dispersibility in PPCC test, and WBL-LMOM do not converge to the results of other statistical models even though 3-hour MPM comparison of 100-hour cumulative data is not an absolute indicator. Therefore, the 3-hour MPM results of WBL-LMOM are relatively unreliable compared to the results of LL-MOM. In the case of GP-MOM, the dispersibility of 3-hour MPM is the second smallest after that of LL-MOM. However, PPCC test shows smaller dispersibility than LL-MOM and converges with 3-hour MPM results of 100-hour cumulative data of other statistical models.



(a) 0.95H



(b) 0.50H



(c) 0.20H

Figure 8. Variance of 3-hour MPM

5 CONCLUSIONS

In this study, the POT method was applied to the statistical model for sloshing impact pressure analysis, and the results according to the threshold ratio were observed for a 1/50 scale 140k LNG cargo hold. Based on the impact pressures obtained from the test, the PPCC test and the 3-hour maximum estimated pressure were compared. The following conclusions can be drawn.

There is a difference in convergence between the statistical models depending on the shape of the excess probability distribution of the experimental results.

As a result of the comparison of 3-hour maximum estimated pressure, overall, the dispersibility tends to decrease as the threshold ratio increases. In the case of the Generalized Pareto distribution, additional consideration is required for the 8% threshold ratio used previously.

As the filling level decreases, the instability of the free surface increases and the sloshing shock pattern becomes more diverse, thus increasing the dispersion of the 3-hour maximum estimated pressure.

In the case of LL-MOM and GP-MOM in this experiment, the dispersibility of 3-hour MPM is low and the results of 100-hour cumulative impact pressure data show convergence with other statistical models.

6 ACKNOWLEDGEMENTS

This study was made possible through the research projects supported by the Lloyd's Register Foundation (LRF) Funded Research Center at Seoul National University (SNU). Its support is greatly appreciated. In addition, the administrative support by the Advanced Marine Engineering Center (AMEC) and the Research Institute of Marine Systems Engineering (RIMSE) of SNU should also be credited.

7 REFERENCES

1. Mathiesen, J. (1976). Sloshing Loads Due to Random Pitching. *Norweg Marit Res*, 3(2), 2-13.
2. Gran, S. (1981). Statistical Distributions of Local Impact Pressures. *Norweg Marit Res*, 8(2), 2-13.
3. Grazcyk, M. Moan, T. Rognebakke, O.(2006). Probabilistic Analysis of Characteristic Pressure for LNG Tanks. *Journal of Offshore Mechanical Architects*, 128, 133-134.
4. Kuo, JF. Campbell, RB. Ding, Z. (2009). LNG Tank Sloshing Assessment Methodology – The New Generation. *Int J Offshore and Polar Eng*, 19(4), 241-253.
5. Kim, SY. Kim, Y. Kim, KH. (2014). Statistical Analysis of Sloshing-Induced Random Impact Pressures, *Proc Inst Mech Eng, Part M: J Eng Maritime Environ*, 228(3), 235-248.
6. Fillon, B. Diebold, L. Henry, J. (2011). Statistical Post-Processing of Long-Duration Sloshing Test. *21st International Ocean and Polar Engineering Conference*, Hawaii, USA.
7. Cetin, E.C. Kim, SY. Kim, Y. (2017). Analysis of Sloshing Impact Pressures Using Different Extreme Statistical Theories. *27th International Ocean and Polar Engineering Conference*, San Francisco, USA.
8. Hosking, JRM. "L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics". In: *Journal of Royal Stat Society Series (B) Method*, 52(1) (1990), 105-124.
9. Heo, J-H, Kho, YW, Shin, H, Kim, S. (2008). Regression Equations of Probability Plot Coefficient test Statistics from Several Probability Distributions. *Journal of Hydrology*, 355, 1-15.
10. Cunnane, C. (1978). Unbiased Plotting Positions – A Review. *Journal of Hydrology*, 37(3/4) ,205-222.
11. Pickands, J III. (1975). Statistical Inference Using Extreme Order Statistics. *Ann Stat*, 3(1), 119-131.