Statistical Post-Processing of Long-Duration Sloshing Test

B. Fillon, L. Diebold, J. Henry, Q. Derbanne, E. Baudin, G. Parmentier
Research Department - Marine Division - Bureau Veritas
Neuilly-Sur-Seine Cdx, France

ABSTRACT

The aim of this study is to compare different statistical fitting distributions of measured sloshing pressures and deduce the best one after having performed for one sea state a 480 hour (at full scale) duration sloshing test (generated by 96 five hour individual tests) at a low partial filling.

KEY WORDS: Sloshing; Sloshing Test; Partial Fillings; Extreme Value Statistics; Offshore LNG Terminals.

INTRODUCTION

Sloshing model tests (submitted by the designer) are standard part of Bureau Veritas comprehensive sloshing assessment (Diebold, 2010). These sloshing model tests determine the sloshing loads to be applied on the Cargo Containment System (hereafter called CCS).

Both observations of the few sloshing events which occurred at sea and sloshing model tests clearly indicate variability of sloshing pressures (Gervaise, De Seze & Maillard, 2009). This stochastic behavior of sloshing pressures result in a flat tail exceeding probability curve. As a consequence, a small change in the probability level (i.e return period) can have strong influence on the statistical pressure. This is the reason why long term approach is considered for the sloshing impact design load assessment.

Indeed, the basic idea of the long term approach is to associate to each one of the sailing condition the ship will face during its lifetime the two following characteristics. First, a short term exceeding probability function of sloshing impact pressure is associated to each sailing condition. Second, a probability of occurrence of impact is associated to each sailing condition. Then the contribution of all the sailing conditions can be cumulated together which results in a long term exceeding probability function (Diebold, 2010). Given this long term exceeding probability function for a specified scenario and cargo containment system capacities (usually normally distributed), the probability of the applied pressures exceeding the structural capacities can be calculated.

Besides the scaling of the model tests sloshing peak pressures (not discussed in this paper), one of the stumbling block of this load assessment is the proper derivation of the short term exceeding probability function for each sailing condition for long return periods as required by the long term approach (Kuo, Campbell & al, 2009). Thus, different statistical fitting distributions (Diebold, 2010) can lead to very different results in terms of design pressures when considering long return periods.

This is the reason why, a dedicated long duration sloshing-test at low partial filling (480 hours at full scale generated by running 96 five hours individual tests) campaign was carried out in cooperation with Ecole Centrale de Nantes. Indeed, this long duration test allowed us to compare the following fitting statistical distributions: Weibull distribution (usually used in industrial practice), the Generalized Pareto Distribution and the Generalized Extreme Value Distribution. Comparisons of these different statistical fitting distributions will be presented through Kolmogorov Smirnov tests, expected pressures for different return periods and confidence intervals.

The conclusion of this study is that the Generalized Pareto Distribution is best suited than the Weibull distribution (generally considered in industrial practice) as sloshing pressure exceeding probability function for the long duration test considered in this study.

SLOSHING TEST

Model Test Setup

The test rig used for the model tests is a six-degree-of-freedom platform called hexapod. This model is a mistral type from SYMETRIE. The specifications of this rig allow us to generate motion within the 1dof limits are given in the following table:

<table>
<thead>
<tr>
<th>Table 1. Motion limits of the hexapod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surge</td>
</tr>
<tr>
<td>Pitch</td>
</tr>
<tr>
<td>Sway</td>
</tr>
<tr>
<td>Roll</td>
</tr>
<tr>
<td>Heave</td>
</tr>
<tr>
<td>Yaw</td>
</tr>
</tbody>
</table>

The maximum load to be tested is 1000 kg still weight and 1000daN of dynamic forces in case of moving masses such as for sloshing tests. The dynamic performances of the actuators reveal a maximum velocity range of 600mm/s and the precision is +/- 1mm for the translation motions and +/- 0.1° for the rotations.

Sloshing model test practice is based on the measurement of fluid impact pressure on the tank walls.

The model (courtesy of GTT) corresponds to the tank N°2 of BV reference vessel with standard cargo capacity of 138 000 m3, scaled to 1/70 and made of a 20 mm thick Plexiglas®. Motions are generated using Froude scaling.
Test Case

In this section are detailed the calculations and validations of the model tank motions.
The considered filling ratio is a low partial filling (R=20%H, where H denotes the tank height) since low partial fillings are known to be the most critical ones in terms of sloshing loads (Diebold, 2010). The ballast loading condition was used for the hydrodynamic calculations. The sea state represented by a Jonswap spectrum considered in this study is the following:

- \( H_s = 8 \text{ m} \), \( T_z = 8.5 \text{ s} \), \( \beta = 195^\circ \), \( \gamma = 2 \) and \( N = 3 \)
- \( V = 0 \text{ kt} \)

where \( H_s \) denotes the significant wave height, \( T_z \) the up-crossing wave period, \( \beta \) the relative wave heading between the waves and the vessel, \( \gamma \) the Jonswap parameter, \( N \) the spreading coefficient and \( V \) the forward speed (in our case, no forward speed).

The mesh used for the hydrodynamic calculations is figured hereafter:

Irregular excitations are derived from the seakeeping calculations performed by HydroStar©. The RAOs for the ship motions are combined with 3600 elementary components of the wave spectrum (Jonswap, \( \gamma = 3.0 \)) for which the phase is defined with a random number. The obtained harmonic motions are then recombined in order to get irregular time history of the ship’s motion for each degree of freedom. For each five hours (at full scale) test, different random numbers are generated in order to create different time histories from the same spectrum.

A particular attention was paid to the irregular excitations generation. Indeed for each test, the 6 degrees of freedom spectral moments (\( m_0 \), \( m_1 \) & \( m_2 \)) were checked against the theoretical values and accepted if and only if the relative error between the calculated and theoretical spectral moments did not exceed 5% for the 6 degrees of freedom.

\[
m_n = \int_0^\infty \omega^n |TF(\omega, \theta)|^2 S(\omega)D(\theta)d\omega, n \in \{0,1,2\}
\]  

(1)

Where \( TF(\omega, \theta) \), \( S(\omega) \) and \( D(\theta) \) denote respectively the transfer function, the Jonswap spectrum and the spreading function.

Model Tests Post-Processing

The impact pressures are measured by dynamic ICP® pressure sensors which natural frequency is above 100 kHz. Static pressure is not taken into account. A total of 54 points around the model could be used to locate the pressure sensors. These sensors are used with dedicated conditioning system composed of ICP cards MEIRI ME26©. The amplified signal is then acquired through an acquisition board National Instruments® NI – PCI6259. A total of 16 pressure sensors is used for the tests. The pressure sensors map used for all the tests is figured hereafter (Fig.3). The sensors indexes are presented in red. Since spreading was considered in the hydrodynamic calculations, the flow at \( R=20\%H \) (where \( H \) denotes the tank height) is essentially transverse even though \( \beta=195^\circ \) was considered. This is the reason why the 4 pressure sensors (Ch. 9, 10, 13 & 14) are highlighted in the graph below since they recorded the highest pressures.

Irregular excitations are derived from the seakeeping calculations performed by HydroStar©. The RAOs for the ship motions are combined with 3600 elementary components of the wave spectrum (Jonswap, \( \gamma = 3.0 \)) for which the phase is defined with a random number. The obtained harmonic motions are then recombined in order to get irregular time history of the ship’s motion for each degree of freedom. For each five hours (at full scale) test, different random numbers are generated in order to create different time histories from the same spectrum.

A particular attention was paid to the irregular excitations generation. Indeed for each test, the 6 degrees of freedom spectral moments (\( m_0 \), \( m_1 \) & \( m_2 \)) were checked against the theoretical values and accepted if and only if the relative error between the calculated and theoretical spectral moments did not exceed 5% for the 6 degrees of freedom.

The sampling rate used is 20 KHz on each channel. The acquisition control program has been developed under DasyLab® by Bureau Veritas.

During the test a pre-analysis is carried out in real time, enabling the operator to assess for each channel the maximum pressure value, the mean value and the number of impacts. This analysis is based on the definition of a time window and a pressure threshold which characterize an event and thus an impact pressure extracted within the trigger condition as the maximum amplitude of this event. The parameters used for the tests hereafter presented are a 400 ms time window and a 20 mbar threshold.

Each channel is processed separately in real time. Thus, each of the 16 pre-processed signals is recorded in a separate file on a hard disk drive (7200 RPM - 16 Mo buffer size). In the same time (buffer mode), all data are saved raw. Hence, any numerical processing applied on site such as triggering, counting and classification methods and other filtering are not modifying the original file. The post-processing of the results is based on the 16 peak files (1 per channel, as far as trigger condition are fulfilled) generated during a test. First, several VBA® routines developed by Bureau Veritas are launched to:
(i) Extract for each channel, elementary statistics such as:
- $P_{\text{max}}$: maximum of impact pressure
- $10P_{\text{max}}$: mean of the 10 higher impacts
- $P_{1/10}$: mean of the tenth of the higher impacts
- $P_{1/3}$: mean of the third of the higher impacts
- $N$: number of impacts

(ii) Build graphic comparisons between selected tests.

Finally, statistical values of impact pressures are computed using dedicated routines developed under statistical softwares (see previous section Test Case), aimed to complete the assessment procedure from the impact pressure point of view.

METHODS FOR STATISTICAL ANALYSIS OF PRESSURE SAMPLES

First of all, the 3 parameters Weibull distribution, noted Weibull3 distribution, has been chosen to fit the pressure samples according to current practice. Then, because fittings obtained by this distribution were not satisfactory, some other approaches have been explored. Especially, the L-moments ratio diagram has been used. This statistical tool provides assistance for the choice of a distribution to fit samples. For the majority of our pressure samples, this diagram gave a Generalized Pareto distribution, noted Gpa distribution, to fit them. The Weibull3 and Gpa distributions, lead to asymptotic distributions for maximum of kind Generalized Extreme Value, noted Gev.

Consequently, these 3 distributions have been studied and their ability and accuracy to estimate sloshing pressures with standard test durations (five hours) have been measured and compared.

Fitting Distributions

Firstly, the pressure samples are fitted by three probability distributions. Their cumulative distribution functions are presented below:

- The Generalized Extreme Value distribution, noted Gev:
  \[ F_X(x) = \exp(-y(x)) \]
  with
  \[ y(x) = \begin{cases} 1 + \frac{\kappa_1}{\kappa_2} \left( \frac{x - \xi_2}{\alpha_2} \right)^{-\frac{1}{\kappa_1}} & \kappa_1 \neq 0 \\ \exp \left( - \frac{x - \xi_2}{\alpha_2} \right) & \kappa_1 = 0 \end{cases} \]
  Where $X$ is a Gev random variable.
  The Gev distribution combines the Gumbel, Fréchet and Reversed Weibull distributions also known as type I, II and III extreme value distributions.

- The Generalized Pareto distribution, noted Gpa:
  \[ F_X(x) = \exp(-y(x)) \]
  with
  \[ y(x) = \begin{cases} \kappa_2^{-1} \left( 1 - \frac{\kappa_2}{\kappa_1} \left( \frac{x - \xi_2}{\alpha_2} \right) \right)^{-\frac{1}{\kappa_2}} & \kappa_2 \neq 0 \\ \left( \frac{x - \xi_2}{\alpha_2} \right) & \kappa_2 = 0 \end{cases} \]
  Where $X$ is a Gpa random variable.
  $\kappa$ is the shape parameter, $\alpha$ the scale parameter and $\xi$ the location parameter. Each parameter is independently estimated for each fitting distribution and each sample.

Fig 4 presents one of our pressure sample of five hours duration fitted by these three distributions. The x-axis is pressure values in mB and the y-axis is the exceedance rate per hour. The exceedance rate is the exceedance probability function multiplied by the events rate. The exceedance probability function (EPF) is
  \[ P[X > x] = \frac{1}{\lambda} \int_x^\infty f(x) \, dx \]
  which is here the probability that a certain pressure value is going to be exceeded. The events rate is the number of events divided by the sample duration (here in hours). Sample is represented by blue dots; fitting by the Weibull3 distribution is represented by a black curve, the one by the Gpa distribution by a red curve and the one by the Gev distribution by a green curve. The 3 and 10 hours return periods are represented with a black horizontal line.

Parameters Estimation for Fitting Distributions

The package “lmomco” in the software R is used for the samples fitting (R Development Core Team, 2010; Asquith, 2009). To estimate the fitting parameters, the package “lmomco” uses a L-moment method.

L-moments are expectations of certain linear combinations of order statistics. They can be defined for any random variable whose mean exists. They suffer less from the effects of sampling variability than conventional moments. This R package uses also the L-moments ratios to estimate parameters of distributions.
Afterwards, parameters of fitting distributions are directly estimated from sample L-moments and sample L-moments ratios.

**Kolmogorov-Smirnov Test for Goodness of Fittings**

Goodness of fitting can be not determinate only graphically, it is important to consolidate our opinion with a statistical test.

The Kolmogorov-Smirnov test is a statistical test for goodness of fittings. It allows knowing if the probability distribution fits well with data.

The statistic of Kolmogorov-Smirnov test is:

\[
\sup_x \left| F_n(x) - F_0(x) \right|
\]

With \( F_n \) the empirical distribution function of data and \( F_0 \) the distribution function used for fitting.

The lower is this statistical value, the better is the fitting. Moreover, this test is used with a rejection criterion corresponding to a probability of 5% to wrongly reject a sample fitting.

Table below shows statistics of Kolmogorov-Smirnov test and the rejection criterion for the 3 fitting distributions and one sample of five hours duration.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Statistic</th>
<th>Rejection criterion</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gev</td>
<td>0.053</td>
<td>( \frac{1.36}{\sqrt{508}} = 0.06 )</td>
<td>Not Rejected</td>
</tr>
<tr>
<td>Gpa</td>
<td>0.039</td>
<td>( \frac{1.36}{\sqrt{508}} = 0.06 )</td>
<td>Not Rejected</td>
</tr>
<tr>
<td>Weibull3</td>
<td>0.073</td>
<td>( \frac{1.36}{\sqrt{508}} = 0.06 )</td>
<td>Rejected</td>
</tr>
</tbody>
</table>

This test only provides criteria. Fitting models which have not been rejected can be considered to be likely close to the actual pressure probability distribution. On the contrary, fitting model which have been rejected are very likely a poor approximation of the actual pressure probability distribution. This means that it is difficult to rely to their pressure estimation for extrapolated return period.

**Confidence Intervals Calculation**

Due to the size limit of the sample, parameters estimation is not exact, but remains random. From designer point of view, punctual estimation is meaningless. A safety margin corresponding to the risk acceptance of the underestimating load pressure is provided by the confidence interval (Fig.8).

Confidence intervals are calculated using a bootstrap percentile method. This method consists to generate \( N \) samples called bootstrap samples, from the initial sample using a random drawing with replacement. Thus, it is possible to obtain similar observations in a bootstrap sample. \( N \) bootstrap samples are the same size as the initial sample. Afterward, each of these \( N \) samples is fitted by a probability distribution. The bounds of the confidence interval are directly obtained from these fittings for the desired confidence level.

**RESULTS OF SAMPLE FITTINGS**

The test of 480 hours duration is built from 96 samples of five hours duration. Each sample of five hours duration is generated with a different seed.

Subsequently, all results are presented for channel 9 which is one of the sensors that recorded highest pressures, unless otherwise noted. This channel 9 is on the tank corner (Fig.3).

**Fittings and Confidence Intervals**

Here, the objective is to check the ability of the 3 stochastic models Gev, Gpa and Weibull3 to accurately estimate, from a sample of limited duration (five hours), pressures associated to greater return periods (like a 400 hours return period).

The quality of fittings is measured from samples of short duration (five hours) taken in the test of 480 hours duration. These samples are fitted and pressures associated to greater return periods than duration of the fitted sample are estimated. Then, these estimations are compared to “reference” values which are the corresponding values of the empirical distribution of the test of 480 hours duration.

Fig.5 presents the full sample of 480 hours duration recorded by channel 9. This is the reference distribution for this channel.

![Fig.5 : The reference distribution of pressures for channel 9.](image-url)
First of all, in order to visualize the differences between the 3 chosen fitting distributions, results of these 3 fitting distributions have been represented on Fig.6 for a sample of five hours duration. It is the same graph that the Fig.4, but fittings are represented until the 400 hours return period.

Sample is represented by blue dots, fitting by the Weibull3 distribution is represented by a black curve, the one by the Gpa distribution by a red curve and the one by the Gev distribution by a green curve.

For estimates associated to return periods which are lower or close to the sample duration, like a 3 hours return period, there is no significant difference between these 3 fitting distributions (Fig.4). However, for estimates associated to return periods which are greater than the test duration, there can be considerable differences between these 3 fitting distributions. For the 400 hours return period, estimate by the Gpa distribution is more than twice greater than those by the Weibull3 distribution; and estimate by the Gev distribution is nearly 5 times greater than the one by the Weibull3 distribution.

Consequently, the choice of fitting distribution is very important and can give very different results.

Fig.7 presents the same results as the Fig.6 but the corresponding reference distribution, represented with purple dots, has been added. Therefore we can directly compare the sample distribution, fitting distributions and the reference distribution.

For this case, the reference distribution is closer to fitting by the Gpa distribution. Nevertheless, results of fittings depend on the sample fitted and so we can obtain very different graphs.

Fig.8 is the same as Fig.6, but confidence intervals for each fitting have been added.

Lower confidence limits are represented with dashed lines and higher confidence limits with dotted lines. Confidence limits and fitting by a Weibull3 distribution are represented in black, confidence limits and fitting by a Gpa distribution in red and confidence limits and fitting by a Gev distribution in green.
Estimated pressure values for fittings by a Gev distribution are greater and confidence intervals are wider than fittings by Gpa or Weibull3 distributions, especially than fittings by a Weibull3 distribution.

The same procedure was performed for each of the 96 pressure samples and each of the 4 channels (channels 9, 10, 13 and 14).

**Kolmogorov Smirnov Tests**

We evaluate goodness of fittings using Kolmogorov Smirnov tests. For each test, the Kolmogorov Smirnov statistic is computed. If this statistic is smaller than an acceptability criterion, the hypothesis that sample fitting is good can not be rejected.

Fig.9 show percentage of acceptable cases on the 96 tests for the 3 fitting distributions for 4 channels: 9, 10, 13 and 14. These sensors give the highest pressures. Results for samples fittings by a Gev distribution are represented in blue, by a Weibull3 distribution in red and by a Gpa distribution in white.

Results of Kolmogorov Smirnov test are better and more homogeneous for fittings by Gpa distribution than for the other distributions.

**Comparison between the Distribution of the 96 Estimated Pressures Associated to One Return Period and the Corresponding Reference Value**

Here, for some return periods, estimated pressure values obtained by fittings are compared to the corresponding reference return period values. The reference return period value means the corresponding pressure value associated to the same return period of sample of 480 hours duration.

Results of 3 and 400 hours return periods for the 3 fitting distributions are presented hereafter. The orange line represents the pressure value associated to a return period of the test of 480 hours duration and black dots represent estimated pressure values associated to the same return period for each of the 96 tests of five hours duration. Purple and green dots represent lower and higher values of confidence intervals at the 95% confidence level. The x-axis represents the index of the 96 tests in a range of 1 to 96.

These graphs are not on the same scale and should be carefully compared.

Pressure values are very often overestimated by the Gev distribution especially for the 400 hours return period.

Fig.10 presents the same graphs for estimated pressure values by a Gpa distribution.

Estimated pressure values by a Gpa distribution seems to better cover pressure values of the test of 480 hours duration. Most of the time, lower and higher values of confidence intervals are on both sides of the orange line.

Fig.12 presents the same graphs for estimated pressure values by a Weibull3 distribution.

Contrary to estimated pressure values by a Gev distribution, the Weibull3 distribution tends to underestimate pressure values.
Percentage of Reference Values Included in the Corresponding Confidence Intervals

For the 96 tests of five hours duration and two return periods (3 and 400 hours), the histograms below present the percentage of confidence intervals that contains the reference pressure values.

Results for sample fittings by a Gev distribution are represented in blue, by a Weibull3 distribution in red and by a Gpa distribution in white.

Fig.13: Percentage of the 96 confidence intervals that contains reference pressure values associated to a 3 and a 400 hours return periods.

Percentages of pressure value included in confidence intervals are more homogeneous for a Gpa distribution. Confidence intervals computed by fittings by Gpa distribution are significantly better than the 2 others fitting distributions, but their size is very large (Fig.11).

Survey of Samples Durations

The previous analysis is done again, by splitting the test of 480 hours duration into 32 tests of 15 hours duration, 16 tests of 30 hours duration, 10 tests of 45 hours duration and 9 tests of 50 hours duration.

For each of the 3 fitting distributions and for the channel 9, Fig.14 compares estimated values for these test durations.

Pressures associated to one return period (3 and 400 hours) are represented by blue dots, lower values of confidence intervals by purple dots and higher values by green dots. Reference values are represented by a black line.

Fig.14: Estimate values for some test durations and the corresponding reference value for channel 9.

When increasing tests duration, estimate values converge, in particular the width of confidence intervals decreases. However, the reference value is not very often included on confidence intervals estimated by the Gev and Weibull3 distributions. Concerning the Gev distribution, estimate values are considerably greater than the reference value; and concerning the Weibull3 distribution, estimate values are very often under the corresponding reference value. The reference value is more often included on confidence intervals estimated by the Gpa distribution. But, for a 400 hours return period, estimate values tend to deviate from the reference value.

It is necessary to increase the test duration. However, it is important to be careful in choosing the fitting distribution to not deviate from the reference value. Even by increasing test duration, differences between the reference value and estimates remain important.
CONCLUSIONS

Statistical post processing of the long duration sloshing test have been carried out on low filling condition for a quasi longitudinal sea heading close to resonance conditions.

The analysis based on this long duration test has shown that the Weibull3 probability model does not allow estimating safely the sloshing pressure on a sensor for a return period greater than the measurement duration.

Gev distribution probability model tends to overestimate the pressure. Gpa distribution probability model seems to be generally consistent with statistical tests for goodness of fit.

Generally, as the scatter of the pressure estimation may be huge particularly when extrapolating on several decades the return period, punctual estimation is not sufficient for design purpose. Confidence limit provides a tool to define a design load with a suitable safety margin.

To avoid over scantling of structural element submitted to sloshing efforts, it may be recommended to design test duration in order to reduce the design pressure confidence limit particularly for critical conditions.

ACKNOWLEDGEMENTS

The sloshing model test campaign was performed in Ecole Centrale de Nantes (ECN).

REFERENCES


