ITTC Quality System Manual

Recommended Procedures and Guidelines

Procedure

Single Significant Amplitude and Confidence Intervals for Stochastic Processes

7.5	Process Control
7.5-02	Testing and Extrapolation Methods
7.5-02-01	General
7.5-02-01-08	Single Significant Amplitude and Confidence Intervals for Stochastic Processes

Updated / Edited by	Approved
Specialist Committee on Stability in Waves of the 28th ITTC	28th ITTC 2017
Date: 04/2017	Date: 09/2017

Table of Contents

Single Significant Amplitude and Confidence Intervals for Stochastic Processes

1. PURPOSE OF PROCEDURE

The purpose of this procedure is to formulate the process for characterizing the uncertainty for data resulting from stochastic processes, such as ship-motion data collected at either model or full scale. Ship-motion data collected in both controlled and uncontrolled environments—typically in a controlled environment for model-scale data and in an uncontrolled environment for full-scale data. The uncertainty of the results includes the following factors:

- Statistical uncertainty caused by finite size of the sample
- Uncertainty in the mass properties of the model or ship
- Uncertainty in significant wave height and peak frequency

Each contribution of uncertainty is characterized by a variance of the estimates.

This procedure only deals with the statistical uncertainty of stochastic data resulting from the finite size of the sample. Although this procedure relates to all stochastic data, data resulting from the response to random excitation, it will be dealt with as though it is ship-motion data.

2. INTRODUCTION

Ship-motion experimental results are characterized in terms of statistical and condition uncertainty. Experimental results are considered random numbers because the environment is intrinsically random and the sample size is finite. Condition uncertainty focuses on environmental wave parameters and the uncertainty related to the mass properties of the model. Condition uncertainty is not treated in this document.

Experimental results are presented in the form of single significant amplitude (SSA), which is a function of the variance estimate of the measured ship motions including surge, sway, heave, roll, pitch, yaw, and lateral and vertical accelerations at various locations on the vessel. Statistical uncertainty will be expressed in terms of confidence intervals for the mean, variance and single significant amplitude (SSA) estimates.

2.1 Statistical Uncertainty

Statistical uncertainty is a result of the finite size of the sample data set, making averages random. The assumption of a normal distribution for these averages is based on the Central Limit Theorem.

A normal distribution is defined by its mean value and variance. The mean value of the estimate approximately equals the estimate itself. The variance of the estimate is computed from the record time-series data. Thus, the uncertainty is quantified by the variance of the estimates (*i.e.*, mean, variance and SSA).

The calculation of the variance of an estimate has to account for the dependency of data points within each record, which are near to each other in time. Two different procedures are provided. One is for a case with a large number of records (runs) while the other is for a smaller number or records (runs).

3. COMPUTING STATISTICAL UNCERTAINTY

The methods used to compute statistical uncertainty will depend on the number of experimental runs and their lengths. If there are a large number of runs, say 30 or more at the same condition, then the methodology is straight forward, and the mean and variance of each run is calculated separately and then they are treated as an ensemble to compute the uncertainty of the runs in total as an ensemble.

In the case of a small number of runs, then there are two more complicated methods for computing the uncertainty, either of which can be applied. However, it must be recognized that these runs must each be suitably long, say on the order of an hour or more, full scale, or the uncertainty can be quite large.

3.1 Large Number of Runs

For cases where data is presented by 30 or more independent records (runs), the following procedure is applicable. The procedure starts by calculating the mean value estimate for each record:

$$\hat{E}_j = \frac{1}{N_j}\sum_{i=1}^{N_j} x_i$$

where x_i is the measured value of the record corresponding to time instant i. Then $\hat{E}_a$, an "ensemble" or "population" mean value estimate for all runs corresponding to a condition, is expressed as:

$$\hat{E}_a = \sum_{j=1}^{Nr} W_j \hat{E}_j \qquad (1)$$

where N_r is the total number of runs for the condition and W_j is the statistical weight of each run, based on the length of each run:

$$W_j = \frac{T_j}{\sum_{i=1}^{Nr} T_i} = \frac{T_j}{T_{Total}} \qquad (2)$$

where T_j is the length of record j and T_{Total} is the cumulative time for all of the records. If all of the records have the same sampling rate (time interval between data points), these weights can be calculated as:

$$W_j = \frac{N_j}{\sum_{i=1}^{Nr} N_i} = \frac{N_j}{N_{total}} \qquad (3)$$

where N_j is the number of data points in each record and N_{Total} is the total number of data points in all records.

The variance of the mean for the ensemble [all records (runs) corresponding to a single condition] can be estimated using the formula:

$$\hat{Var}\left(\hat{E}_a\right) = \sum_{i=1}^{Nr} W_i^2 \left(\hat{E}_i - \hat{E}_a\right)^2 \qquad (4)$$

An estimate of the variance of the variance for the ensemble is calculated starting with estimating the variance for each record of the ensemble. This is calculated using the ensemble mean as:

$$\hat{V}_j = \frac{1}{N_j - 1}\sum_{i=1}^{N_j}\left(x_i - \hat{E}_a\right)^2$$

The next step is to calculate the ensemble-averaged variance:

$$\hat{V}_a = \sum_{j=1}^{Nr} W_j \hat{V}_j \tag{5}$$

Then the variance of the variance estimate is computed as

$$\hat{Var}\left(\hat{V}_a\right) = \sum_{i=1}^{Nr} W_i^2 \left(\hat{V}_i - \hat{V}_a\right)^2 \tag{6}$$

The results of Eqs. (4) and (6) are carried forward for the assessment of uncertainty of the final result. A more detailed discussion of the calculation of the estimates and the derivation of these formulae can be found in Belenky, *et al*. (2013).

3.2 Small Number of Runs

For cases where data is presented by fewer than 30 independent records (runs), the variance of the variance estimate can be evaluated using the estimate of the auto-covariance function of centered squares (Belenky, *et al*. 2015) or by direct counting. In either case, the estimates of the mean value and variance are calculated from the original record using all available data.

3.2.1 Variance of the Variance by Auto-Covariance Function of Centered Squares

The confidence interval for a statistical estimate provides a practical quantification of its statistical uncertainty and is generally calculated from the variance of the statistical estimate. Belenky *et al*. (2015) describes a technique for calculating the variance of the mean and variance of the variance estimates for a sample consisting of a few relatively long records.

The estimates are calculated from an estimate of the auto-covariance function of the process, $\hat{R}(\tau_i)$:

$$\hat{R}(\tau_i) = \frac{1}{N}\sum_{j=1}^{N-i}(x_j - \hat{E})(x_{j+i} - \hat{E}) \ ,$$

where τ_i is the *i*th time lag, $\hat{E}$ is the estimate of the mean value, and N is total number of data points available; and from an estimate of the auto-covariance function of the centered squares of the process, $\hat{R}_y(\tau_i)$:

$$\hat{R}_y(\tau_i) = \frac{1}{N}\sum_{j=1}^{N-i}\left(\left(x_j - \hat{E}\right)^2 - \hat{V}\right)\left(\left(x_{j+i} - \hat{E}\right)^2 - \hat{V}\right)$$
$$= \frac{1}{N}\sum_{j=1}^{N-i}\left(y_j - \hat{V}\right)\left(y_{j+i} - \hat{V}\right)$$

where $y_j = \left(x_j - \hat{E}\right)^2$ and $\hat{V}$ is the estimate of the variance of the process.

Using the auto-covariance functions of the process and of the centered squares of the process, the variances of the estimates of a single record are computed as:

$$Var\left(\hat{E}\right) = \frac{\hat{V}}{N} + \frac{2}{N}\cdot\sum_{i=1}^{M-1}\left(1 - \frac{i}{M}\right)\cdot\hat{R}(\tau_i) \tag{7}$$

$$Var\left(\hat{V}\right) = \frac{\hat{V}_y}{N} + \frac{2}{N}\cdot\sum_{i=1}^{M-1}\left(1 - \frac{i}{M}\right)\cdot\hat{R}_y(\tau_i) \tag{8}$$

where $\hat{V}_y$ is an estimate of the variance of the centered squares and M is the number points for which both auto-covariance functions are estimated. The value of M is chosen in the interval:

$$0.5\sqrt{N} \le M \le 2\sqrt{N} \tag{9}$$

This recommendation for *M* is based on Priestley (1981); Kiefer & Vogelsgang (2002); Brockwell & Davis (2006); and Sun (2014). Additional sources include Crowson (1963a, 1963b), which point out the necessity for a dramatic decrease in *M* relative to *N*.

Consider a sample case consisting of a single, very long sample record: Run A. Estimates of the auto-covariance function of the process and of its central squares are shown in Figures 1 and 2, respectively.

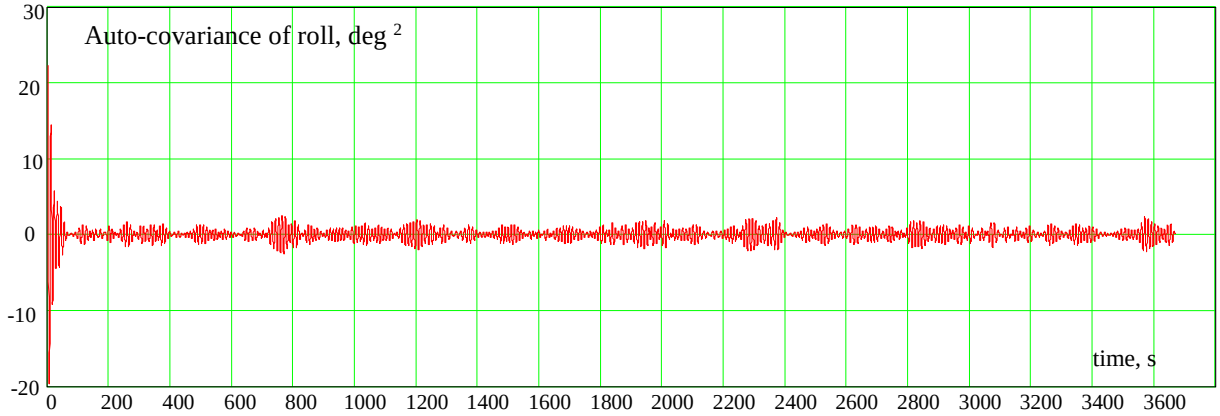

Figure 1 — Estimate of the Auto-Covariance Function of Roll Channel, Run A

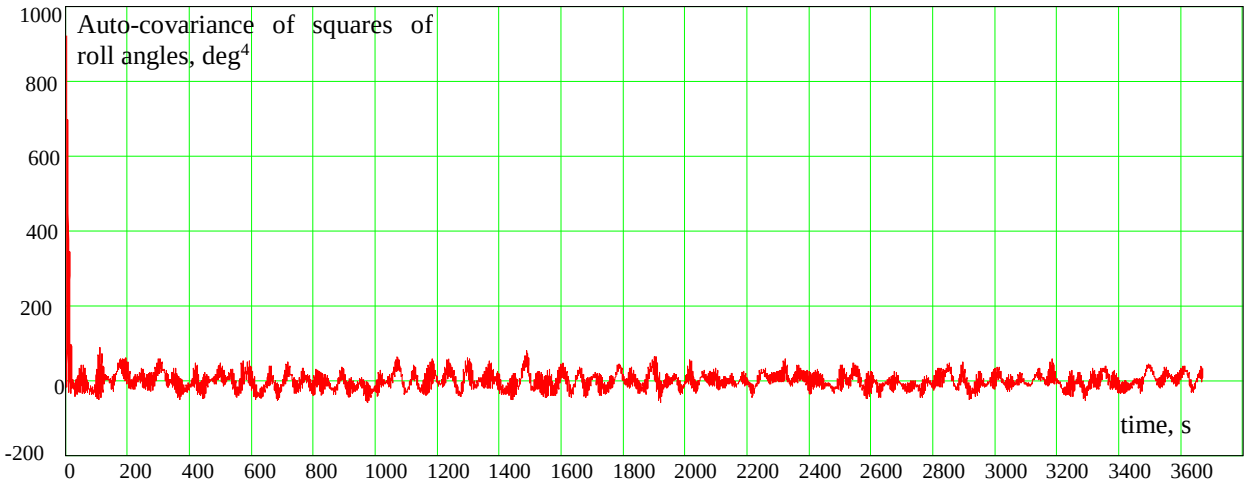

Figure 2 — Estimate of the Auto-Covariance Function of Squared Values - Roll Channel, Run A

Table 1 — Comparison of Different Options for *M*

Run	Value	$M = 0.5\sqrt{N}$ Eqs. (7) & (8)	$M = \sqrt{N}$ Eqs. (7) & (8)	$M = 2\sqrt{N}$ Eqs. (7) & (8)
A	Var. of Mean	0.001493	0.0008543	0.0005382
	Var. of Var.	0.7764	0.9358	0.9900
	Mean Upper	-0.3218	-0.3402	-0.3521
	Mean Lower	-0.4733	-0.4548	-0.4430
	SSA Upper	9.8120	9.8478	9.8592
	SSA Lower	9.0807	9.0449	9.0335
B	Var. of Mean	0.007340	0.005063	0.005162
	Var. of Var.	7.7460	8.9473	9.2669
	Mean Upper	3.5685	3.5400	3.541
	Mean Lower	3.2326	3.2611	3.260
	SSA Upper	17.203	17.253	17.265
	SSA Lower	15.884	15.835	15.823

The accuracy of the estimate of the auto-covariance function deteriorates with the increase of the time lag, because the volume of available data is decreased. In order to deal with deteriorated accuracy of the auto-covariance function for large lags, a linear weighting function and averaging over even a few records, described in Belenky *et al*. (2013), has been found to give reasonable results. However, the case considered consists of only one record. As a result, the numerical error remains in the estimate, reaching approximately 7 to 8-percent of the first term.

Comparisons of the results for two sample runs, Runs A and B using Eqs. (7) and (8) is given in Table 1.

The boundaries of SSA show insignificant influence of choice of *M* using Eq. (9); the difference can be seen only in the second digit after the decimal point.

3.2.2 Variance of the Variance by Direct Counting

The objective is to find a statistical estimate of SSA and evaluate the uncertainty of this estimate. The sample is presented as a set of records of a stationary process; lengths of the records may differ from each other.

For a normal (Gaussian) process, the 95.45th-percentile (often just called the 95th percentile) corresponds to the average of the one-third largest amplitudes; this is defined as the Single Significant Amplitude (SSA). In direct counting, the average of the one-third largest amplitudes will be used to compute the SSA, independent of any assumptions as to the distribution of the records being analysed.

The statistical estimation of SSA via direct counting is free from any assumptions on the distribution of the process $x(t)$. However, it is still assumed that the process is stationary. The sample of the process is represented as a set of independent records of different lengths, while

dependence may be significant within each record. A nested array (array consisting of records of different length) is a convenient way to describe this type of data:

$$X = \left\{ x_j \right\}_i ; \quad j = 1,...,N_i ; \quad i = 1,...,Nr$$

where X is used to identify the entire sample, the index i, identifies a record and the index j a point within a record. Nr is the number of records available in the sample, and N_i is the number of data points in a specific record.

First, the mean value of the process $x(t)$ is estimated based on the sample X:

$$\hat{E}_x = \sum_{i=1}^{Nr} W_i \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \left\{ x_j \right\}_i \right)$$

where W_i, the weighting factors for each record, calculated in according to Eq. (2) or (3). Next, the mean-crossing points (times) are found. These points are also presented as a nested array:

$$TM = \left\{ Tm_j \right\}_i ; \quad j = 1,...,Nm_i ; \quad i = 1,...,Nr$$

where Nm_i is the number of mean crossings in the ith record.

The next step is to search for the peak values, which are defined by the largest absolute value of the sample between each pair of consecutive crossings of the level of the mean value estimate, see Figure 3. The total number of peak values will be:

$$Np = \sum_{i=1}^{Nr} (Nm_i - 1)$$

The peak data must include the time corresponding to these points. The data may be presented in the form of a matrix with three columns—record index, time of peak value in the record, and peak value—and Np rows. The matrix is defined as follows:

$$Pk = \begin{cases} Pk_{k,1} = i; \\ Pk_{k,2} = t_j; \\ \quad \text{if } \left\{ x\left(t_j\right) \right\}_i = \max \left\{ x\left(t_j\right) \right\}_i \\ \qquad \cap\, t_j \in \left[\left\{ Tm_j \right\}, \left\{ Tm_{j+1} \right\} \right] \\ Pk_{k,3} = \left\{ x\left(t_j\right) \right\}_i \end{cases}$$

$$j = 1,...,N_i ; \quad i = 1,...,Nr ; \quad k = 1,...,Np,$$

where k is the index of the peak value, which is advanced every time a mean-crossing peak is identified.

The next step is to find the level corresponding to the one-third largest peak values. It is, in fact, an estimation of the 1/3rd quantile.

It is found by sorting the peaks by value (highest to lowest) and finding the value of the peak that encompasses the highest of 1/3rd set of the sorted list of peaks:

$$Pks = \text{sort}\left(Pk^{<3>} \right)$$

$$ids = \text{round}\left(\frac{Np}{3} \right)$$

$$\hat{a}_{1/3} = Pks_{ids}$$

here $Pk^{<3>}$ means the "3rd column of the matrix Pk".

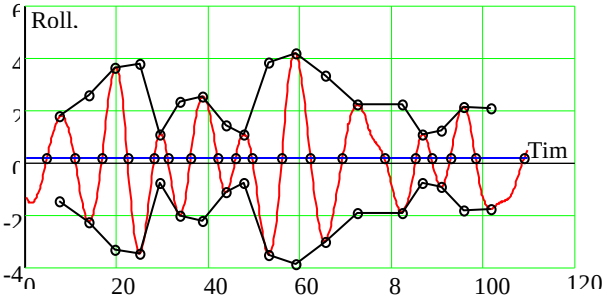

Figure 3 — Definition of Mean Crossing Peaks

It is found by sorting the peaks by value (highest to lowest) and finding the value of the peak that encompasses the highest of 1/3rd set of the sorted list of peaks:

$$Pks = \text{sort}\left(Pk^{<3>}\right)$$
$$ids = \text{round}\left(\frac{Np}{3}\right)$$
$$\hat{a}_{1/3} = Pks_{ids}$$

here $Pk^{<3>}$ means the "3rd column of the matrix Pk".

The sample of the 1/3rd largest peaks is then extracted from the matrix Pk:

$$Ps = \begin{cases} Ps_{l,1} = Pk_{k,1}\,; \\ Ps_{l,2} = Pk_{k,2}; \quad if \quad Pk_{k,3} \geq a_{1/3} \\ Ps_{l,3} = Pk_{k,3} \end{cases}$$
$$k = 1,..,Np \quad ; \quad l = 1,..,Ns$$

where Ns is the total number peaks above the 1/3rd-quantile estimate and will be, by definition, the same as the index ids above:

$$Ns = ids = \text{round}\left(\frac{Np}{3}\right)$$

The estimate of SSA is the estimate of the mean value of $Pk^{<3>}$:

$$\hat{\text{SSA}} = \hat{\text{E}}(Ps^{<3>}) = \frac{1}{Ns}\sum_{l=1}^{Ns} Ps_{l,3} \qquad (10)$$

Figure 4 illustrates the peak average calculation for a sample roll time history. Peak values are marked in green or blue. The level of the 1/3rd highest peak values is set as the 66.6th percentile of all peaks and is shown as the green line. The peaks above this value are marked in blue and comprise the 1/3rd largest peaks. The average of these blue peaks is the SSA estimate and is shown as a blue line.

3.2.3 Uncertainty of Statistical Estimation of SSA

The statistical estimate of SSA by direct counting, Eq. (10) is a random number, so a confidence interval is needed to assess the statistical uncertainty of the estimate. As the estimation of this mean involves the summation of instances of random variables, it is reasonable to assume a normal distribution for the estimate. To define the normal distribution, one needs to compute the mean value and the variance of the estimate.

As the mean is known to be an unbiased estimate, the mean of the estimate is equal to the estimate itself:

$$\text{E}(\hat{\text{SSA}}) = \text{E}\left(\frac{1}{Ns}\sum_{l=1}^{Ns} Ps_{l,3}\right) = \hat{\text{SSA}} \qquad (11)$$

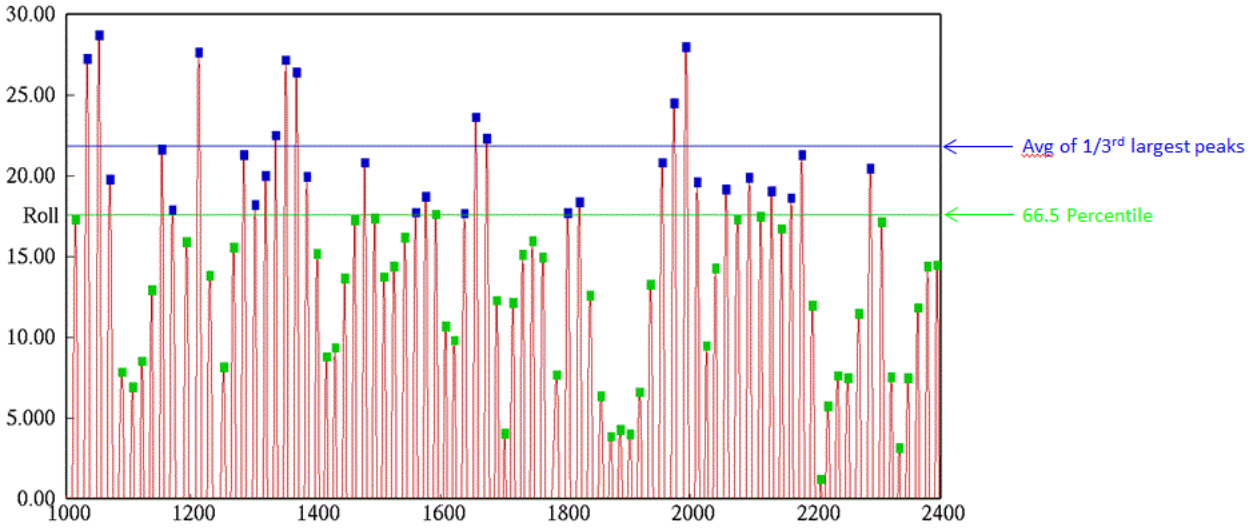

Figure 4 — Peak Average Calculation

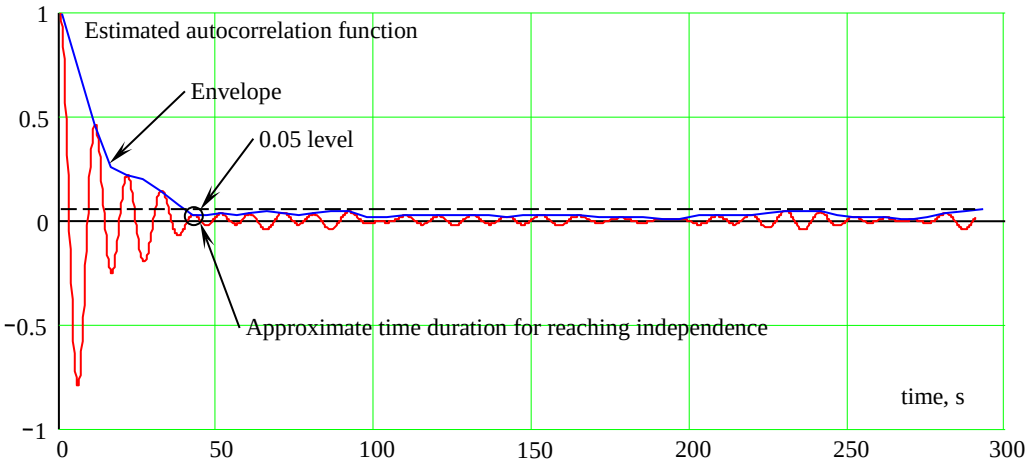

Figure 5 — Estimated Auto-correlation for Roll Motion

However, in order to find the variance of the estimate, Eq. (10), the dependence within the sample $Ps^{<3>}$ must be addressed. The 1/3rd-largest peaks, as well as all of the peaks, are the values of the motion process at particular instants of time, which may be dependent on one another if they are from the same record and are close together in time.

Statistical dependence within the ship-motion process is a reflection of the physical nature of the mechanical motion of a rigid body on the surface of a dense fluid. Several phenomena contribute to this dependence. First, waves themselves carry the dependence because water is a dense fluid and motion of water possesses significant inertia. A ship acting essentially as a

filter transforms this dependence in terms of its response. Second, a ship has its own inertia, which also contributes to dependence. Third, there are hydrodynamic forces proportional to the accelerations, generally referred to as added masses, which also are contributing factors. The fourth is the hydrodynamic memory effect, which is a result of the radiation and diffraction of waves. The overall dependence can be characterized through an estimate of the auto-covariance function of the motion data. The typical appearance of the estimate of the autocorrelation (normalized auto-covariance) function for roll motions is shown in Figure 5.

The time difference τ_{ind} for which the motions (or peaks) can be assumed to be independent is approximated by the time lag where the normalized auto-covariance function falls below the value of 0.05 (taken as a significance level). As the auto-correlation function of ship motions is oscillatory, it is convenient to use its envelope instead. The envelope is also plotted in Figure 5 by connecting the absolute value of the peaks of the auto-correlation function by a blue line.

The peaks that were recorded within the time difference τ_{ind}, are assumed to be dependent. These peaks are grouped. These groups comprise a series of successive peaks, which are from the same record and for which the time increment between peaks is less than τ_{ind}. As these groups will consist of different number of peaks, a nested array is a natural form for organizing this data:

$$PS = \left\{ Ps_j \right\}_i ; \quad j = 1,...,Npc_i ; \quad i = 1,...,Nps$$

where Npc_i is the number of peaks in the ith "group" and Nps is the number of "groups" found while analyzing the dependence between the 1/3rd-largest peaks.

Once the peaks have been sorted into these groups, the variance of the peak average estimate can be computed using a procedure similar to that for the mean or variance of a sample consisting of a few, relatively long records (Belenky *et al.* 2013). In that procedure, the data from different records was assumed to be independent, and the estimate of the auto-covariance in time was used to compute the effect of dependence within each record. To maximize accuracy, an average auto-covariance function is computed for all records.

The calculation for the variance of the peak average follows this same approach, except that

- Peaks from different groups are assumed to be independent
- Auto-covariance with respect to the difference in peak index, rather than time lag, is used to account for the dependence between peaks in a group
- Auto-covariance is averaged over groups, rather than records

This averaged auto-covariance function for the 1/3rd-largest peaks is estimated as:

$$\hat{R}s_j = \frac{1}{Ns} \sum_{i=1}^{Nps} \sum_{k=1}^{Npm-j} \begin{cases} \left(\left\{ Ps_k \right\}_i - S\hat{S}A \right) \\ \quad \times \left(\left\{ Ps_{k+j} \right\}_i - S\hat{S}A \right), \\ \quad \text{if} \quad Npc_i \geq k + j \\ 0, \quad \text{if} \quad Npc_i < k + j \end{cases} \qquad (12)$$

where Npm is the maximum number of peaks in a group:

$$Npm = \max\left(Npc_i \right)$$

Figure 6 shows an example of the averaged auto-covariance for roll as estimated using Eq.

(12). As can be seen, the dependence covers three sequential peaks.

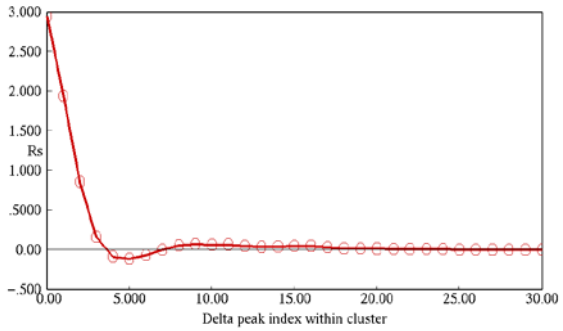

Figure 6 — Estimate of Averaged Auto-Covariance for 1/3rd Largest Peaks for Roll Motion

Once the averaged auto-covariance function has estimated with Eq. (12), the variance of the SSA estimate can be estimated as follows:

$$\mathrm{Var}(\hat{\mathrm{SSA}}) = \frac{Rs_0}{Ns} + \frac{2}{Ns} \times \sum_{i=1}^{Nps} \sum_{k=1}^{M_i - 1} \left(1 - \frac{k}{M_i}\right) Rs_k \frac{Npc_i}{Ns} \tag{13}$$

where M_i is the summation cutoff for each group, which is set to:

$$M_i = \sqrt{Npc_i}$$

In theory, the summation can be carried out to $Npc_i - 1$, but the estimate of the auto-covariance function becomes inaccurate at large separation times (or peak indices in this calculation) as the number of data pairs used in calculating the function becomes small.

Once the variance of estimate has been calculated, the assumption of a normal distribution for the SSA estimate allows of the confidence interval for SSA as:

$$\hat{\mathrm{SSA}}_{\mathrm{Low,Hi}} = \hat{\mathrm{SSA}} \pm \mathrm{K}_\beta \sqrt{\mathrm{Var}(\hat{\mathrm{SSA}})} \tag{14}$$

where K_β is a factor based on the quantile of a normal distribution for a specified level of probability. For a typical probability level of 0.95, K_β is approximately 1.96.

4. CONFIDENCE INTERVAL

4.1 General

The confidence interval (CI) is defined for a sample estimate and reflects its uncertainty, which is caused by its random nature. A two-sided CI is an interval around the estimate where the true value is contained with a given probability (confidence probability or confidence level).

The calculation of the confidence interval requires a distribution of the estimate. The lower and upper boundaries S_{Low} and S_{Up} of the confidence interval of the estimate S are calculated with quantiles (inverse of the cumulative distribution function). For two-sided confidence interval, these boundaries are

$$S_{Low} = Q\left(\frac{1 - P_\beta}{3}\right); \quad S_{Up} = Q\left(\frac{1 + P_\beta}{3}\right) \tag{15}$$

where P_β is the accepted confidence probability and Q stands for quantile.

As stated in Section 3, the normal distribution is assumed for all values except for significant wave height and modal period. For the normal distribution, the center of the range is determined by the calculated statistical estimate. The

width of the range is determined from the variance of this estimate. Note that this does *not* assume that the distribution of the process itself is normal.

If a normal distribution of the estimate is assumed, Eq. (15) can be simplified, because the normal distribution is symmetric

$$S_{Low} = S - K_\beta \sqrt{Var(S)}$$
$$S_{Up} = S + K_\beta \sqrt{Var(S)}$$

where $Var(S)$ is the variance of the estimate S and the coefficient K_β is one-half the non-dimensional width of the confidence interval and is calculated as:

$$K_\beta = Q_N\left(\frac{1+P_\beta}{2}\right)$$

where Q_N is a quantile of the standard normal distribution (zero-mean & unity-variance). For the typical probably of 0.95, K_β equals 1.96

4.2 Confidence Interval for a Mean Value Estimate

The upper and lower bounds of the confidence interval for the mean values are:

$$E_{Low} = \hat{E}_a - K_\beta \sqrt{Var(\hat{E}_a)}$$

$$E_{Up} = \hat{E}_a + K_\beta \sqrt{Var(\hat{E}_a))}$$

In these expressions, $\hat{E}_a$ is the mean estimate for the ensemble, Eq. (1); and $Var(\hat{E}_a)$ is the estimate of the variance of the mean for the ensemble, Eq. (4).

4.3 Confidence Interval for a Variance Estimate

Expressions for the confidence interval of the variance are:

$$V_{Low} = \hat{V}_a - K_\beta \sqrt{Var(\hat{V}_a)}$$

$$V_{Up} = \hat{V}_a + K_\beta \sqrt{Var(\hat{V}_a)}$$

where $\hat{V}_a$ is the variance estimate for the ensemble, calculated with Eq. (5); $Var(\hat{V}_a)$ is the variance of the variance estimate for the ensemble, calculated with Eq. (6).

5. SINGLE SIGNIFICANT AMPLITUDE (SSA)

The following equation relates a generic variance estimate $\hat{V}_X$ for a process $X(t)$ and its SSA:

$$\hat{SSA}_X = 2 \cdot \sqrt{\hat{V}_X} \qquad (16)$$

Strictly speaking, Eq. (16) defines single significant amplitude for a normal process only. If $X(t)$ is normal, SSA_X is the mean value of the 1/3 largest amplitudes. If Eq. (16) is used without assumption of normality $X(t)$, it can be considered as a convenient form of representation for variance and/or standard deviation estimates.

The delta method is used to calculate the variance of the SSA. The upper boundary is determined assuming a normal distribution for the variance of the SSA estimate. The bias introduced by the square root function is assumed to be small, following the conventional assumption of small bias of standard deviation.

Equation (16) is a deterministic function of a single random argument—the variance estimate

$\hat{V}_X$. By the delta method, this function is expanded in a Taylor series around its mean value $V_0 = E(\hat{V}_X)$

$$S\hat{S}A(\hat{V}_X) \approx S\hat{S}A(V_0) + \left.\frac{dS\hat{S}A(\hat{V}_X)}{d\hat{V}_X}\right|_{\hat{V}=\hat{V}_0} \left(\hat{V}_X - V_0\right) + \frac{1}{2}\left.\frac{d^2 S\hat{S}A(\hat{V}_X)}{d\hat{V}_X^2}\right|_{\hat{V}=\hat{V}_0} \left(\hat{V}_X - V_0\right)^2 + \ldots \quad (17)$$

In the delta method, Eq. (17) is linearized, so the terms of second order and higher are neglected. The Evaluation of the derivative and substitution into Eq. (16) yields the following linearized function:

$$S\hat{S}A_{Lin}(\hat{V}_X) = 2\sqrt{V_0} + \frac{\left(\hat{V}_X - V_0\right)}{\sqrt{V_0}} \quad (18)$$

The variance of the variance estimate is known: $Var\left(\hat{V}_X\right)$ for each process of interest. The mean value of the estimate $V_0 = E(\hat{V}_0)$ is a deterministic value, *i.e.*, is a constant in (17).

$$Var(V_0) = 0 \text{ and } Var\left(2 \cdot \sqrt{V_0}\right) = 0$$

The application of the rule for the variance of a linear function, applied to equation (18) yields:

$$Var\left(S\hat{S}A_{Lin}(\hat{V}_X)\right) = \frac{Var\left(\hat{V}_X\right)}{V_0}. \quad (19)$$

The true mean value for the variance estimate $V_0 = E(\hat{V}_X)$ is not known, so it is replaced by its estimate, which is the variance estimate itself:

$$V_0 \approx \hat{V}_X$$

Eq. (19) becomes:

$$Var\left(S\hat{S}A_{Lin}(\hat{V}_X)\right) = \frac{Var\left(\hat{V}_X\right)}{\hat{V}_X}. \quad (20)$$

Finally, the upper and lower boundaries of the SSA are calculated as:

$$SSA_{U,L} = SSA \pm K_\beta \sqrt{Var\left(S\hat{S}A_{Lin}(\hat{V}_X)\right)} \quad (21)$$

6. COMPARISON OF RESULTS

The final result of the procedure is an estimate of the Single Significant Amplitude (SSA) and its confidence interval.

Two different methods for calculating this estimate have been presented:

- Indirect method—using the estimate of the variance
- Direct method—using the direct counting of zero-crossing peaks.

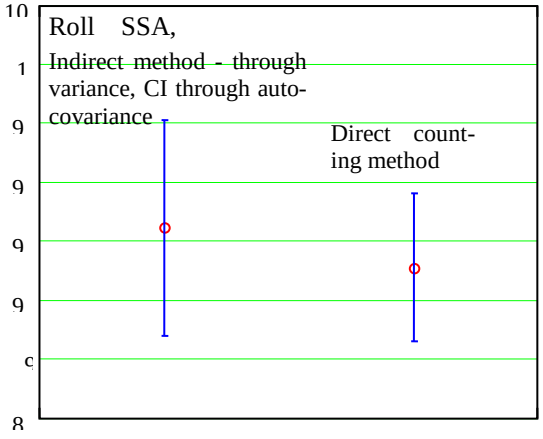

Figure 7 — Comparison of Different Methods of Calculation of SSA with CI, Run A

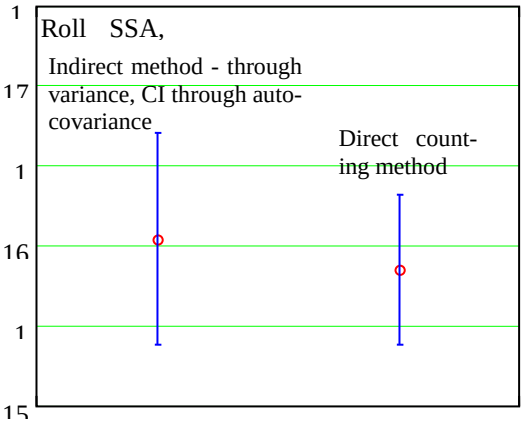

Figure 8 — Comparison of Different Methods of Calculation of SSA with CI, Run B.

Two methods have been presented for assessing the confidence interval:

- Through auto-covariance estimate, using the variance of the variance Eq. (8) and then Eqs. (20) and (21)
- Using the uncertainty of direct counting of zero crossing peaks, Eqs. (13) and (14)

Figures 7 and 8 compare SSA estimates with confidence interval as computed using the direct and indirect methods and with the confidence interval computed with the two methods.

As each estimate is within the confidence interval of the other, the estimates are statistically equivalent.

As each estimate is within the confidence interval of the other, the estimates are statistically equivalent.

Comparison of the two estimates and two methods of evaluation of the confidence interval lead to the conclusions:

- Estimate of SSA calculated from variance is statistically identical to the estimate of SSA calculated by averaging peaks.
- Confidence interval of the SSA calculated by direct counting of peaks results in the smallest value relative to the other methods.

In general, smaller confidence interval means the data is used more effectively, resulting in better accuracy of the estimate. Based on 10 cases, direct counting of peaks is the most efficient approach for processing the data. However, the difference between direct counting and using the variance was small for all 10 cases; so the practical impact is likely to be negligible.

7. REFERENCES

Belenky, V., V. Pipiras, C. Kent, M. Hughes, B. Campbell & T. Smith (2013) "On the Statistcal Uncertainties of Time-domain-based Assessment of Stability Failures: Confidence Interval for the Mean and Variance of a Time Series." *Proc. 13th Int'l. Ship Stability Workshop*, Brest, France, pp. 251–258.

Belenky, V., V. Pipiras, & K. Weems (2015) "Statistical Uncertainty of Ship Motion Data." *Proc. 12th Int'l. Conf. Stability of* Ships *and Ocean Vehicles*, Glasgow, UK, 12 p.

Brockwell, P. J. & R. A. Davis (2006) *Time Series:* Theory *and Methods*. Springer, 2nd edition.

Crowson, H. L. (1963a) "An Error Analysis in the Digital Computation of the Autocorrelation Function." *AIAA J.*, 1(2):488–489.

Crowson, H. L. (1963b) "Generalization of the Note 'An Error Analysis in the Digital Computation of the Autocorrelation Function'." *AIAA J.*, 1(8):1968–1969.

Kiefer, N. M. & T. J. Vogelsgang (2002) "Heteroskedasticity autocorrelation robust testing using bandwidth equal to sample size." *Econometric Theory*, 18:1350–1366.

Priestley, M. B. (1981) *Spectral Analysis and Time* Series. Volume 1, Academic Press.

Sun, Y. (2014) "Lets fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference." *J. Econometrics*, 178:659–677.