Statistical Uncertainty of Ship Motion Data

Vadim Belenky, *NSWCCD (David Taylor Model Basin)*, vadim.belenky@navy.mil

Vladas Pipiras, *University of North Carolina at Chapel Hill*, pipiras@email.unc.edu

Kenneth Weems, *NSWCCD (David Taylor Model Basin)*, kenneth.weems@navy.mil

ABSTRACT

Records of nonlinear ship motion data, which are the basis for a probabilistic assessment of dynamic stability of a ship in irregular waves, are produced by time-domain numerical simulations or model tests in a basin. The volume of such samples is finite, so any statistical estimates calculated from a sample are random numbers and need to have a confidence interval, which quantifies the statistical uncertainty of the estimate. Ship motion data samples generally come in the form of an ensemble of records for a given condition, in which dependence may be very strong within the record, while the records themselves are independent of one another.

Since multiple data points describe the same feature of the process, statistically dependent data usually contains less information in comparison to independent data, so the confidence interval is wider for a set of dependent data than for the same amount of independent data. The paper revisits known mathematical methods to account for data dependency in computing the variance of the mean estimate and the variance of the variance estimate, which are the basis for computing a confidence interval of these estimates. The paper also addresses the calculation of the variance of the mean and variance of the variance for an ensemble of independent records of different length. The issue of minimum record length is considered and it is shown that a record of any length can contribute to the ensemble estimates of mean and variance.

Keywords: *Confidence interval, Statistical estimate*

1. INTRODUCTION

The development of probabilistic models for the assessment of the dynamic stability of a ship requires a characterization of the nonlinear response of the ship to severe sea conditions. This characterization is generally based on time-domain numerical simulations or model-scale experiments in large, "random" waves derived from theoretical or experimental representations of severe ocean waves. The direct results of such model test or numerical simulation campaigns are presented as a set of time histories of ship motion in large amplitude, irregular waves.

As the waves are irregular, the time histories are records of a stochastic process of ship motions. The most basic statistical processing includes the estimation of the mean value and variance or standard deviation. These estimates are essentially random numbers, which tend to the true value as the volume of data increases. A confidence interval is a measure of how close the estimate is likely to be to the true value, and is presented as a range of values and a probability, P_{β}, that the true value is within that range. For example, an estimate of standard deviation with its confidence interval can be presented as:

$$\hat{\sigma} = [\sigma_{low}; \sigma_{up} \mid P_{\beta}] \qquad (1)$$

It means that there is a P_β chance that the true value of the standard deviation is between the lower boundary σ_{low} and upper boundary σ_{up}. The confidence probability P_β is an *a priori* given or agreed value; P_β=0.95 is widely used for engineering purposes. The "hat" above a symbol indicates an "estimate", which is a random value. The absence of a "hat" indicates that the value is deterministic.

Engineering calculations will typically use the upper and/or lower values of the confidence interval as a bound of the actual expected value. For example, if a measure of the intensity of ship motion is needed for further assessment or calculation, the upper boundary of confidence interval can be used as a conservative value. A change to the confidence probability P_β allows the conservatism of the assessment to be "tuned" to a level appropriate for the task in hand. A larger P_β will result in a wider confidence interval for an estimate and a wider range of values for assessments based on that estimate, but reduce the likelihood that the true value of the assessment is outside of the computed range.

Confidence intervals are heavily used in validation as they enable the comparison of two estimates; this application aspect (among others) is considered in Smith and Zuzick (2015) and is outside of the scope of this paper. Another use for the confidence interval is in the planning of model tests and numerical simulations, as it can help to determine the number and length of model tests or simulation runs that are needed to achieve required accuracy; this type of application is also outside the scope of this paper.

The uncertainty of statistical estimates is of particular concern for assessments, which involve the prediction of extreme responses or low-probability events form non-linear time domain data. Since these assessments are fundamentally extrapolations, the uncertainty in the results will tend to be very sensitive to the uncertainty in the statistical characterization of the data. For this reason, the consideration of the statistical uncertainty of ship motions is an important part of the ONR project "A Probabilistic Procedure for Evaluating the Dynamic Stability and Capsizing of Naval Vessels" (Belenky, *et al.*, 2015).

2. CONCEPT OF ENSEMBLE

The ensemble is a set of ship motion data records which represent a single or narrow range of sea and operating conditions. By its definition, it presumes that more than one record may be needed. Why is this so?

- Limitations on record length; for model tests in a seakeeping basin, the limited size of the facility will limit the duration of any test run with forward speed. As a result, a single record may have too few wave encounters to assess motions. This will particularly be true for cases with high speed and/or following or quartering seas.

- Practical non-ergodicity; the nonlinearity of ship behavior may cause one run to be insufficient for a complete assessment, even if it is relatively long. A typical example is parametric roll in head seas (Reed 2011), for which a typical run in a linear basin did not provide the necessary variation in initial conditions for proper statistical characterization of parametric roll.

- Valid modeling of irregular waves for numerical simulation. The elevation, pressure and velocity field of the incident wave is generally modeled using Fourier series, where amplitudes are defined by spectra and phases are random. The duration for which such a model will produce statistically independent waves will depend on the number of frequencies used for the discretization of a spectrum (Belenky, 2011). Increasing the number of frequencies in the wave model incurs a significant computational cost, so a set of relatively short records, each of which requires fewer wave frequencies, is computationally more efficient than one long record requiring many frequencies.

The records in ensemble are not necessarily of the same length. It is both difficult and unnecessary to ensure that experimental runs have exactly the same duration.

In the analysis of irregular wave motion data, the processes of waves and ship motions are assumed to be stationary. If the ship capsizes, the response of the "mast-down" ship will be fundamentally different for its upright response, which has to be considered as a violation of the stationarity of the motion process. Attempting to include pre-capsize, capsize and post-capsize motion as part of a single stationary process will make the required volume of the data impractical. As a result, the record has to be cut immediately prior to capsizing. Similarly, it may be necessary to truncate a model test run if variations in the speed or relative heading of the ship become too large. Both of these scenarios may result in records of different lengths.

Thus, the ensemble for a particular wave environment, speed and heading is an irregular data structure that can be described by a "nested array" defined as an array that contains other arrays as elements. To avoid confusion with typical matrix notation, the following nomenclature will be used:

$$X = [x_i]_j \; ; \quad i = 1,...Np_j \; ; \quad j = 1,..Nr \qquad (2)$$

The index within the square bracket refers to a data point within a record while the index outside of the square brackets relates to the record number. An example data structure is illustrated in Figure 1.

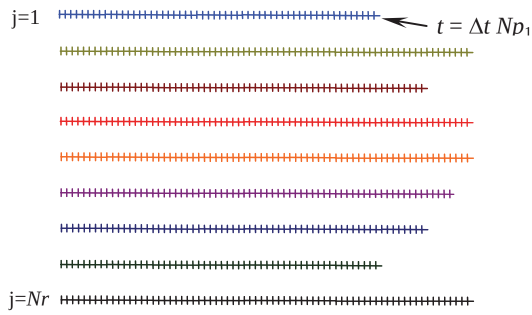

Figure 1: Illustration of a nested array

The data points within each record are dependent, while no dependence in expected between the records. Dependence between the data points is a result of inertia of ship motions, hydrodynamic memory and inertia of water particles in wave. Independence between simulation records is ensured by using different pseudo-random sets of initial phases in the model of the irregular waves. Independence between model test records is supported by the pseudo-random actuation of the wave maker and by the time lag between runs; this time usually is sufficient for the waves to be radiated and decay on the damping beach of a basin.

The combination of dependent and independent data within a single sample is specific to ship motion data.

3. ESTIMATES FOR A SINGLE RECORD

The consideration starts with examining mean value and variance estimates and their confidence interval for a single record. To simplify the notation, the square brackets and record index are not used in this section.

3.1 Dependence and Uncertainty

Consider the mean value and variance estimates.

$$\hat{E} = \frac{1}{Np} \sum_{i=1}^{Np} x_i \qquad (3)$$

$$\hat{V} = \frac{1}{Np-1}\sum_{i=1}^{Np}\left(x_i - \hat{E}\right)^2 \qquad (4)$$

The dependency within a record does not matter for these estimates. Changing the sampling rate or time increment will change the number of data points, but it will be reflected in the number *Np*. As long as the time increment remains within the reasonable range, *e.g.* it does not become large compared to the response period of the ship, the estimates (3) and (4) are affected very little.

This is not true for the confidence interval. The confidence interval is a metric of statistical uncertainty, which generally decreases with an increased volume of the sample. Theoretically, the width of the confidence interval goes to zero when the volume of the sample becomes infinite, because the estimate becomes a true value. Increasing the sampling rate does not increase the amount of information available in the sample; however the number of data points becomes larger. The data points become closer to each other. Since the same information is carried by more points, the dependence between data points becomes stronger, and the contribution of each of them is decreased.

Conversely, if the increment between the data points is increased, their dependence is decreased and the contribution of each individual data point becomes larger. Further increase of the increment (decrease of the sampling rate) should lead to independence. Once the independence is achieved, the contribution of each data point can no longer be affected by other points. This means that the number of independent points will define the amount of information available in the sample. The dependence between the data points may therefore have a serious effect on uncertainty and the width of the confidence interval. The mathematical treatment of this influence is considered further.

3.2 Variance of the Mean Value

Estimates of the mean value and variance of the process *X* are random numbers and, as any other random numbers, may have a variance. Priestley (1981) gives a general direction for the derivation of the formulae for the variance of mean value and variance estimates. That derivation was reproduced in Belenky *et al.* (2013) in order to examine the role of the assumption of normal distribution for *X*. An abridged version of this derivation in included below for the sake of completeness.

Apply a variance operator to both sides of equation (3) and treat the sum as if the values are dependent, so the variance of the sum is a sum of all of the terms of the covariance matrix:

$$\begin{gathered}\mathrm{var}\left(\hat{E}\right) = \mathrm{var}\left(\frac{1}{Np}\sum_{i=1}^{Np} x_i\right) = \\ \frac{1}{Np^2}\sum_{i=1}^{Np}\sum_{j=1}^{Np} Cov(x_i, x_j)\end{gathered} \qquad (5)$$

Var(..) is the variance operator and *Cov*(..) is the covariance operator. Since the process *X* is assumed to be stationary, its auto-covariance function depends only on the difference in time (time lag) between the two points and does not depend on particular time instances:

$$\begin{gathered}Cov(x_i, x_j) = R(t_{i-j}) = R(\tau_k) \\ k = 0, 1..., Np-1\end{gathered} \qquad (6)$$

Consider a sum of all the elements of the covariance matrix in Equation (5):

$$\sum_{i=1}^{Np}\sum_{j=1}^{Np} Cov(x_i, x_j) = \qquad (7)$$

$$\sum\begin{pmatrix} R(\tau_0) & R(\tau_1) & \ldots & R(\tau_{N-2}) & R(\tau_{N-1}) \\ R(\tau_1) & R(\tau_0) & \ldots & R(\tau_{N-3}) & R(\tau_{N-2}) \\ \ldots & \ldots & \ldots & \ldots & .. \\ R(\tau_{N-2}) & R(\tau_{N-3}) & \ldots & R(\tau_0) & R(\tau_1) \\ R(\tau_{N-1}) & R(\tau_{N-2}) & \ldots & R(\tau_1) & R(\tau_0) \end{pmatrix}$$

The elements of the main diagonal are variances:

$$R(\tau_0) = R(0) = V \qquad (8)$$

The other elements on the line parallel to the main diagonal are also the same; the next element to the term $R(\tau_0)=V$ is always $R(\tau_1)$, then $R(\tau_2)$ and so forth. The main diagonal of a $Np\times Np$ square matrix contains Np elements. The lines of elements parallel to the main diagonal and located next to it contain only Np-1 elements. Each subsequent line will have one fewer element, until diagonals at the low-left or upper-right corner have only one element. Having in mind that the covariance matrix is symmetric relative to its main diagonal and all the "lines of elements" except the main diagonal are encountered twice:

$$\begin{aligned}\sum_{i=1}^{Np}\sum_{j=1}^{Np} Cov(x_i, x_j) &= Np\cdot V + \\ &2\big((Np-1)R(\tau_1) + (N-2)R(\tau_2) + ... + R(\tau_{N-1})\big) = \\ &Np\cdot V + 2\cdot\sum_{i=1}^{Np-1}(N-i)R(\tau_i)\end{aligned} \qquad (9)$$

Substitution of Equation (9) into Equation (5) leads to the standard formula for the variance of the mean value estimate (see *e.g.* Priestly 1981):

$$\mathrm{var}\big(\hat{E}\big) = \frac{V}{Np} + \frac{2}{Np}\cdot\sum_{i=1}^{Np-1}\left(1-\frac{i}{Np}\right)\cdot R(\tau_i) \qquad (10)$$

The first term in Equation (10) is actually the variance of the mean estimate of a random variable, while the second term accounts for the dependence between the data points of a stochastic process. As expected, if the process X is uncorrelated white noise (Wiener process), the result is identical to the one for a random variable, because the auto-covariance function of the white noise equals zero for all non-zero time lags.

3.3 Variance of the Variance

Variance is, by definition, the average of centered squares, so a process Y is introduced as:

$$y_i = (x_i - E)^2 \approx \big(x_i - \hat{E}\big)^2 \qquad (11)$$

The estimate of the mean value of the process Y is the estimate of the variance of the original process x:

$$\hat{E}_y = \hat{V} \qquad (12)$$

The variance of the mean estimate of the process Y is then the variance of the variance estimate of the process X:

$$\mathrm{var}\big(\hat{V}\big) = \frac{V_y}{Np} + \frac{2}{Np}\cdot\sum_{i=1}^{Np-1}\left(1-\frac{i}{Np}\right)\cdot R_y(\tau_i) \qquad (13)$$

V_y and R_y are, respectively, the variance and the auto-covariance function of the process of centered squares Y.

The standard formula for the variance of the variance (*e.g.* Priestley 1981) uses the assumption that the process X is normal, which leads to

$$\mathrm{var}\big(\hat{V}\big) = \frac{2V^2}{Np} + \frac{4}{Np}\sum_{i=1}^{Np-1}\left(1-\frac{i}{Np}\right)\big(R(\tau_i)\big)^2 \qquad (14)$$

Because for the normal process

$$V_y = 2\cdot V^2\,; \quad R_y(\tau) = 2\cdot\big(R(\tau)\big)^2 \qquad (15)$$

Reed (2011) uses an alterative form of (14):

$$\mathrm{var}\big(\hat{V}\big) = \frac{2}{Np}\sum_{i=-(Np-1)}^{Np-1}\left(1-\frac{|i|}{Np}\right)\big(R(\tau_{|i|})\big)^2 \qquad (16)$$

As noted in Belenky, *et al*. (2013), there is no apparent reason to use the normal assumption for the process X. The calculation of the auto-covariance function of the centered squares requires little additional computation effort in comparison with the straight auto-covariance function.

3.4 Estimate of Auto-Covariance

To use Equations (10) and (13), it is necessary to estimate the auto-covariance functions of the processes *X* and *Y*. The estimate is expressed as:

$$\hat{R}'(\tau_i) = \frac{1}{Np - i}\sum_{j=1}^{Np-i}(x_j - \hat{E})(x_{j+i} - \hat{E}) \qquad (17)$$

Accuracy of the estimate (17) deteriorates very quickly for larger time lags due to insufficient data – as the time lag gets larger there are fewer pairs of data points with that time lag. This leads to statistical "noise" as shown Figure 2. This is obviously noise as there is no reason why the dependence could that strong after 500 seconds.

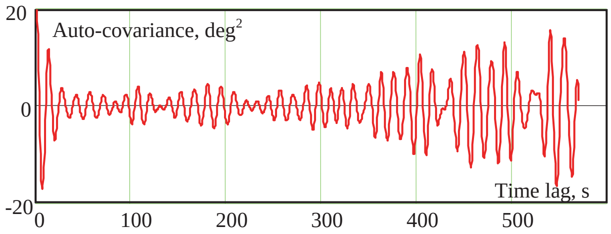

Figure 2 Estimate of auto-covariance function

This loss of accuracy can be alleviated by a simple weighting factor: (*Np-i*)/*Np*. Such weighting results in little change to the auto-covariance function for small time lags as the difference between *Np* and *Np-i* is not significant for small *i*. When the index *i* becomes large, the amount of available data decreases and therefore the influence of its contribution also decreases. The weighted estimate is expressed as:

$$\hat{R}(\tau_i) = \frac{1}{Np}\sum_{j=1}^{Np-i}(x_j - \hat{E})(x_{j+i} - \hat{E}) \qquad (18)$$

The result of weighting the estimate of the auto-covariance function is shown in Figure 3. It is apparent that the amount of "noise" has subsided, while the initial part (first 100 seconds) has not changed very much. Details on the numerical example can be found in Belenky *et al.* (2013).

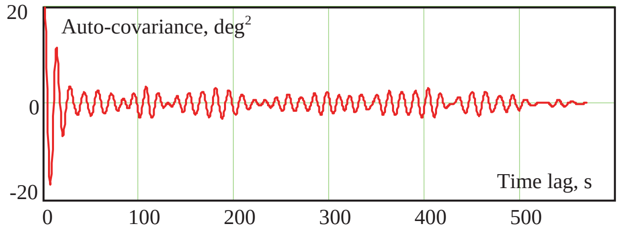

Figure 3 Weighted estimate of auto-covariance function

However, weighting the estimate may not be sufficient to get rid of all of the "noise". Cases are still possible when the "noise" makes the calculations completely senseless (*e.g.* producing negative value of the variance of the mean) if one uses the estimate (18) in formulae (10) or (13). Since the auto-covariance estimates at large lags are still not very reliable, they can be cut off, at a point designated *M*. Equations (10) and (13) are re-written as:

$$\hat{\text{var}}(\hat{E}) = \frac{V}{Np} + \frac{2}{Np}\cdot\sum_{i=1}^{M-1}\left(1 - \frac{i}{M}\right)\cdot\hat{R}(\tau_i) \qquad (19)$$

$$\hat{\text{var}}(\hat{V}) = \frac{V_y}{Np} + \frac{2}{Np}\cdot\sum_{i=1}^{M-1}\left(1 - \frac{i}{M}\right)\cdot\hat{R}_y(\tau_i) \qquad (20)$$

Belenky, *et al.* (2013) considered *M*=*Np*/2, which works well if the estimate of the auto-covariance is fairly accurate. Further review of the literature has led to (Priestly, 1981; Brockwell and Davis, 2008):

$$M = \sqrt{Np} \qquad (21)$$

Some sources also suggest $2\sqrt{Np}$ or $\sqrt{Np}/2$. The origin of this formula is optimality of spectral smoothing. The range [$0.5Np^{0.5}$; $2Np^{0.5}$] appears to represent an area where the result is not very sensitive to the specific value of *M*. The operation of cutting off the autocorrelation function is essentially the same as smoothing the spectral estimate. Spectral representations are a traditional way of processing ship motion information and can also be used for the estimate of the auto-covariance. However, the discussion of estimation of spectra is outside of the scope of this paper.

4. ESTIMATES OF AN ENSAMBLE

4.1 Estimate of Mean and Variance

Consider an ensemble of Nr records, each of which has Np_j data points. The time increment Δt is assumed to be the same for all the records, which is the usual practice for both numerical simulations and model tests. The statistical weight for each record is expressed as follows:

$$W_j = \frac{Np_j}{Nt} \tag{22}$$

Nt is the total number of points in the ensemble:

$$Nt = \sum_{j=1}^{Nr} Np_j \tag{23}$$

The ensemble estimate for the mean value is calculated for all of the points:

$$\hat{E}_a = \frac{1}{Nt}\sum_{j=1}^{Nr}\sum_{i=1}^{Np_j}[x_i]_j = \frac{1}{Nt}\sum_{j=1}^{Nr}\frac{Np_j}{Np_j}\sum_{i=1}^{Np_j}[x_i]_j = \sum_{i=1}^{Nr} W_i\left(\frac{1}{Np_j}\sum_{i=1}^{Np_j}[x_i]_j\right) = \sum_{i=1}^{Nr} W_i \hat{E}_i \tag{24}$$

$\hat{E}_j$ is the mean value estimate for a record j. The ensemble estimate of the variance is:

$$\hat{V}_a = \frac{1}{Nt-1}\sum_{j=1}^{Nr}\sum_{i=1}^{Np_j}\left([x_i]_j - \hat{E}_a\right)^2 = \sum_{i=1}^{Nr} W_j{}'\hat{V}_j{}' \tag{25}$$

$$\hat{V}_j{}' = \frac{1}{Np_j - 1}\sum_{i=1}^{Np_j}\left([x_i]_j - \hat{E}_a\right)^2 \tag{26}$$

$$W_j{}' = \frac{Np_j - 1}{Nt - 1} \tag{27}$$

The weights (27) are slightly different from (22). However, as the number of points is quite large (thousands and tens of thousands), one can state that

$$W_j \approx W_j{}' \tag{28}$$

Note that the variance estimate in (26) is not exactly the same as the record variance estimate from (4), as it uses the ensemble mean estimate instead of record mean estimate.

4.2 Estimate of Auto-Covariance Function

As the records may have different length, the estimate of the auto-covariance function (18) is padded with zeros to facilitate averaging across the record:

$$[\hat{R}_m]_j = \frac{1}{Np_j}\begin{cases}\sum_{i=1}^{Np_j - m}([x_i]_j - \hat{E}_a)([x_{i+m}]_j - \hat{E}_a) & i+m < Np_j \\ 0 & i+m \geq Np_j\end{cases} \tag{29}$$

$$m = 1,.., \max\left(Np_j\right)$$

Like the data, the record estimate of the auto-covariance is presented in a form of nested array, with j being the index of record, while m is the index of the time lag. Since they have been padded by zeros, all of the record estimates of the auto-covariance function have the same length.

The ensemble estimate of the auto-covariance function is obtained by averaging across the records (assuming that if very short records are present in the ensemble, their statistical weight is small):

$$\hat{R}_a(\tau_m) = \sum_{j=1}^{Nr} W_j[\hat{R}_m]_j = \frac{1}{Nt}\sum_{j=1}^{Nr}\sum_{i=1}^{Np_j - m}\begin{cases}([x_i]_j - \hat{E}_a)([x_{i+m}]_j - \hat{E}_a) & i+m < Np_j \\ 0 & i+m \geq Np_j\end{cases} \tag{30}$$

Note that for m=0, equation (30) yields an expression identical to the formula for ensemble averaged variance. The averaging procedure significantly decreases the amount of “noise”, as illustrated in Figure 4.

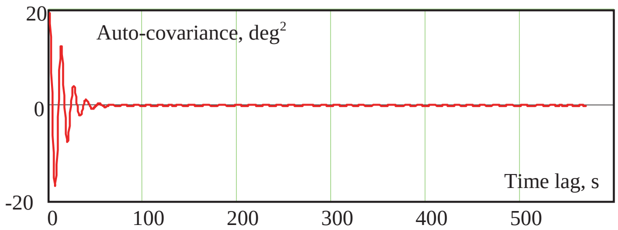

Figure 4 Ensemble-averaged estimate of auto-covariance function (Belenky *et al.*, 2013)

The formula for the ensemble averaged estimate for the process Y (the process of the centered squares) is similar to equation (30):

$$\hat{R}_{Ya}(\tau_m) = \frac{1}{Nt}\sum_{j=1}^{Nr}\sum_{i=1}^{Np_j-m}\begin{cases}([y_i]_j - \hat{V}_a)([y_{i+m}]_j - \hat{V}_a) & i+m < Np_j \\ 0 & i+m \geq Np_j\end{cases} \quad (31)$$

The ensemble-averaged estimate of auto-covariance function of centered squared for an example set of roll data is plotted in Figure 5.

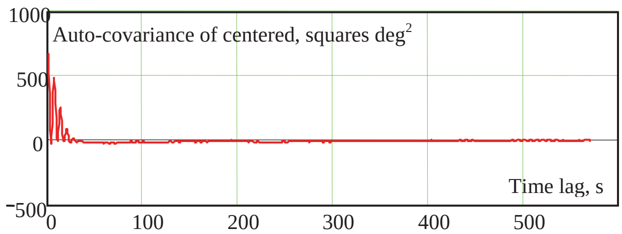

Figure 5: Ensemble-averaged estimate of auto-covariance function of centered squares (Belenky *et al.*, 2013)

4.3 Variances of Mean and Variance

In order to get the variance of the ensemble-averaged mean estimate, the variance operator is applied to both sides of equation (24):

$$\mathrm{var}(\hat{E}_a) = \mathrm{var}\left(\sum_{j=1}^{Nr} W_j \hat{E}_j\right) = \sum_{j=1}^{Nr} W_j^2\, \mathrm{var}(\hat{E}_j) \quad (32)$$

$\mathrm{var}(\hat{E}_j)$ is the variance of the record mean value estimate, expressed with equation (19) where the auto-covariance function is estimated by equation (29). The cut-off point M can be taken for the ensemble:

$$M = \sqrt{\max(Np_j)} \quad (33)$$

Substitution of equations (19), (22) and (29) into (32) leads to:

$$\hat{\mathrm{var}}(\hat{E}_a) = \frac{1}{Nt}\sum_{j=1}^{Nr} W_j \hat{V}_j' + \frac{2}{Nt}\sum_{m=1}^{M}\left(1-\frac{m}{M}\right)\cdot\sum_{j=1}^{Nr} W_j [\hat{R}_m]_j \quad (34)$$

Here, the variance estimate (26) is used instead of the record estimate (4) for consistency with the auto-covariance estimate (29), so the ensemble mean estimate is used instead of record mean estimate.

Equations (25), (28) and (30) can be used to re-write equation (34) in terms of ensemble-averaged estimates:

$$\hat{\mathrm{var}}(\hat{E}_a) = \frac{\hat{V}_a}{Nt} + \frac{2}{Nt}\sum_{m=1}^{M}\left(1-\frac{m}{M}\right)\cdot\hat{R}_a(\tau_m) \quad (35)$$

A similar argument can be made for the ensemble-averaged variance of the variance estimate:

$$\hat{\mathrm{var}}(\hat{V}_a) = \sum_{i=1}^{Nr} (W_j')^2\, \mathrm{var}(V_j) \quad (36)$$

$$\hat{\mathrm{var}}(\hat{V}_a) = \frac{\hat{V}_{Ya}}{Nt} + \frac{2}{Nt}\sum_{m=1}^{M}\left(1-\frac{m}{M}\right)\cdot\hat{R}_{Ya}(\tau_m) \quad (37)$$

$\hat{V}_{Ya}$ is the ensemble-averaged estimate of the variance estimate of the process Y (centered squares) based on the ensemble-averaged mean value estimate:

$$\hat{V}_{Ya} = \hat{R}_{Ya}(0) = \frac{1}{Nt}\sum_{j=1}^{Nr}\sum_{i=1}^{Np_j}([y_i]_j - \hat{V}_a)^2 \quad (38)$$

4.4 Alternative Method for Variances of Mean and Variance

If the number of records is large enough, the variances of the mean and variance estimates can be computed without an auto-covariance estimate. Consider equation (32) for the special case where all of the records have the same length, so all the weights are the same and equal to $1/Nr$. The theoretical values of the variances of the mean estimates $\mathrm{var}\left(\hat{E}_j\right)$ are the same for all the records. The variance of ensemble-averaged mean estimate for the records of the same length is expressed as:

$$\mathrm{var}\left(\hat{E}_a\right) = \frac{\mathrm{var}\left(\hat{E}_j\right)}{Nr} \qquad (39)$$

If the bias is assumed to be small, the estimate can be used instead of theoretical value:

$$\hat{\mathrm{var}}\left(\hat{E}_a\right) = \frac{1}{Nr^2}\sum_{j=1}^{Nr}\left(\hat{E}_i - \hat{E}_a\right)^2 \qquad (40)$$

Equation (40) can then be presented as:

$$\hat{\mathrm{var}}\left(\hat{E}_a\right) = \sum_{j=1}^{Nr} W^2\left(\hat{E}_i - \hat{E}_a\right)^2 \qquad W = \frac{1}{Nr} \qquad (41)$$

The weight lacks a record index as all of the records are of the same length. However, this requirement is no longer necessary as the weight is inside the summation sign, so the record index j can be brought back:

$$\hat{\mathrm{var}}\left(\hat{E}_a\right) = \sum_{j=1}^{Nr} W_j^2\left(\hat{E}_i - \hat{E}_a\right)^2 \qquad (42)$$

Formula (42) is equivalent to formula (35). To prove this, start by substituting (3) into (39):

$$\hat{\mathrm{var}}\left(\hat{E}_a\right) = \sum_{j=1}^{Nr} W_j^2\left(\frac{1}{Np_j}\sum_{i=1}^{Np_j}[x_i]_j - \hat{E}_a\right)^2 = \sum_{j=1}^{Nr} W_j^2\left(\frac{1}{Np_j}\sum_{i=1}^{Np_j}[x_i]_j - \frac{1}{Np_j}\sum_{i=1}^{Np_j}\hat{E}_a\right)^2 = \sum_{j=1}^{Nr}\frac{W_j^2}{Np_j^2}\left(\sum_{i=1}^{Np_j}\left([x_i]_j - \hat{E}_a\right)\right)^2 \qquad (43)$$

Using the well-known formula for the square of a sum:

$$\left(\sum_{i=1}^{N} a_i\right)^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j = \sum_{i=1}^{N} a_i^2 + \sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N} a_i a_j = \sum_{i=1}^{N} a_i^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{N-i} a_i a_{i+j} \qquad (44)$$

Note the algebraic equivalence of the structure of equations (44) and (7). Applying the expansion (44) to equation (43) leads to:

$$\hat{\mathrm{var}}\left(\hat{E}_a\right) = \sum_{j=1}^{Nr}\frac{W_j^2}{Np_j^2}\left(\sum_{i=1}^{Np_j}\left([x_i]_j - \hat{E}_a\right)^2 + 2\sum_{m=1}^{Np_j}\sum_{i=1}^{Np_j-m}\left([x_i]_j - \hat{E}_a\right)\left([x_{i+m}]_j - \hat{E}_a\right)\right) \qquad (45)$$

The first term of (45) can be converted into a biased estimate of the record estimate and the second term can be converted into a non-weighted estimate of the auto-covariance function (17):

$$\hat{\mathrm{var}}\left(\hat{E}_a\right) = \sum_{j=1}^{Nr}\frac{W_j^2}{Np_j}\left(V_j'' + 2\sum_{i=1}^{Np_j}\frac{Np_j - i}{Np_j}\cdot\frac{1}{Np_j - i}\sum_{k=1}^{Np_j-i}\left([x_i]_j - \hat{E}_a\right)\times\left([x_{i+k}]_j - \hat{E}_a\right)\right) = \sum_{j=1}^{Nr} W_j^2\left(\frac{\hat{V}_j''}{Np_j} + \frac{1}{Np_j}\sum_{m=1}^{Np_j}\left(1 - \frac{m}{Np_j}\right)\cdot[\hat{R}_m']_j\right) \qquad (46)$$

$\hat{V}_j$" is the biased estimate of record variance.

$$\hat{V}_j" = \frac{Np_j - 1}{Np_j} \hat{V}_j' \approx \hat{V}_j' \qquad (47)$$

$[\hat{R}_m']_j =$

$$\frac{1}{Np_j - m} \begin{cases} \sum_{i=1}^{Np_j - m} ([x_i]_j - \hat{E}_a)([x_{i+m}]_j - \hat{E}_a) \\ \qquad i + m < Np_j \\ 0 \\ \qquad i + m \geq Np_j \end{cases} \qquad (48)$$

$m = 1,.., \max(Np_j)$

Substitution of equation (22) into (46) yields

$$\hat{\text{var}}(\hat{E}_a) = \frac{1}{Nt} \sum_{j=1}^{Nr} W_j \hat{V}_j" + \frac{1}{Nt} \sum_{j=1}^{Nr} W_j \sum_{m=1}^{Np_j} \left(1 - \frac{m}{Np_j}\right) \cdot [\hat{R}_m']_j \qquad (49)$$

Introducing the cut-off point M defined by equation (33) completes the derivation

$$\hat{\text{var}}(\hat{E}_a) = \frac{\hat{V}_a"}{Nt} + \frac{1}{Nt} \sum_{m=1}^{M} \left(1 - \frac{m}{M}\right) \cdot \hat{R}_a'(\tau_m) \qquad (50)$$

Equation (50) is identical to equation (35), taking into account that the large number of points and insignificant bias:

$$\hat{V}_a" = \sum_{j=1}^{Nr} W_j \hat{V}_j" \approx \hat{V}_a \qquad (51)$$

Estimates of auto-covariance (29) and (48) differ by weighting. However, they can still be considered to be approximately equal because the cut-off limits the influence of weighting, so

$$\hat{R}_a'(\tau_m) = \sum_{j=1}^{Nr} W_j [\hat{R}_m']_j \approx \hat{R}_a(\tau_m) \qquad (52)$$

A similar argument can be made for the variance of the variance, allowing the following formula to be used for the calculation of the variance of the ensemble-averaged variance estimate:

$$V\hat{a}r(\hat{V}_a) = \sum_{j=1}^{Nr} W_j^2 (\hat{V}'_i - \hat{V}_a)^2 \qquad (53)$$

Further consideration of equation (53) can be found in Belenky *et al*., (2013).

4.5 Confidence Interval for Mean and Variance Estimates

The calculation of the boundaries of the confidence interval requires knowledge of the distribution of the estimates. This information is rarely available as distribution of the estimate is related with the distribution of the process itself. For example, if a sample of independent random variables is known to have normal distribution, the estimate of the mean will have student-t distribution and the distribution of variance estimate is related to 2 distribution.

The distributions of the processes of ship motion are not known. Even if the central part of the distribution can be approximated with normal for some motions and some ships, the mutual dependence of data points creates difficulties with using Student-t and 2 distribution. On the other hand, the sample, i.e. ensemble of records, is presented with large number of points. The calculations of the estimates involve mostly summation, so it seems appropriate to invoke the Central Limit Theorem, which allows the distribution of the estimates to be assumed to be normal.

This assumption presents no difficulties for the mean value, but may be a problem with variance estimate. The normal distribution supports negative values, while the variance and its estimate cannot be negative. Practical experience, however, shows that the confidence interval of variance is usually small enough to keep the low boundary of the variance far from zero. Nevertheless, the possibility of numerical difficulties does exist, especially for smaller ensemble data volume.

Once the assumption of normality of distribution of the estimate is accepted, the calculation of the boundaries of confidence interval is trivial:

$$E_{low} = \hat{E}_a - K_\beta \sqrt{Var(\hat{E}_a)}$$
$$E_{up} = \hat{E}_a + K_\beta \sqrt{Var(\hat{E}_a)} \tag{50}$$

$$V_{low} = \hat{V}_a - K_\beta \sqrt{Var(\hat{V}_a)}$$
$$V_{up} = \hat{V}_a + K_\beta \sqrt{Var(\hat{V}_a)} \tag{51}$$

K_β is the $0.5(1+P_\beta)$ quantile of a standard normal distribution (with zero mean and unity variance):

$$K_\beta(P_\beta) = Q_N\left(\frac{1+P_\beta}{2}\right)$$
$$K_\beta(0.95) = 1.959964.. \approx 1.96 \tag{52}$$

The confidence interval for standard deviation can be calculated using the "boundary" method (Bickel and Doksum, 2001):

$$\sigma_{low} = \sqrt{V_{low}} \quad ; \quad \sigma_{up} = \sqrt{V_{up}} \tag{53}$$

5. CONCLUSIONS AND FUTURE WORK

The analysis of dynamic stability in ocean waves is based primarily on irregular sea ship motion data obtained from model tests in the basin or time-domain nonlinear numerical simulations. As the volume of data from these sources is, by necessity, limited, such analyses must account for uncertainties that result from the finite volume of data. The present paper presents robust and easy-to-use formulae for the calculation of estimates of the mean value and variance, with confidence intervals, from such data.

Ship motions in irregular waves are generally presented as an "ensemble" of records of time-domain data which has been computed or measured for the same environmental conditions, loading conditions, speed and heading. The records are independent of each other, but there is a strong dependence between data points within each record. Different records may have different length, so the natural data structure for an ensemble is a nested array (*i.e.* an array containing other arrays).

The structure of the dependence (strong dependence within each record and independence of records to each other) does not affect the ensemble-averaged estimates of mean value and variance, but must be accounted for when evaluating the statistical uncertainty of those estimates. The dependence within each record is accounted for through estimates of the auto-covariance function of the value of ship motion processes and their centered squares. As these quantities are estimated from a finite-length time series, a cut-off point is introduced to limit the possible influence of statistical "noise" caused by a deterioration of accuracy for large time lags. The estimation of the auto-covariance functions may be avoided if an ensemble contains a sufficient number of independent records.

Future development may be expected in the relation of the statistical uncertainty with spectral characteristics. In particular, the smoothed spectral estimate can be seen as a natural source for the estimate of auto-covariance function. Further test calculations are desirable in order to determine how many independent records are "sufficient" to use formulae (39) and (48) instead of (32) and (37).

Future work may also include further testing of the formulae. This would include creating or collecting a large set of ensembles from different experimental and numerical sources in order to see how well the computed confidence interval captures the expected values of the ensemble estimates. The fraction of estimates falling within the confidence interval should be close to the given confidence probability.

6. ACKNOWLEDGEMENTS

The work described in this paper has been funded by the Office of Naval Research, under Dr. Patrick Purtell, Dr. Ki-Han Kim and Dr. Thomas Fu. The authors greatly appreciate their support. The participation of Prof. Pipiras was facilitated by the Summer Faculty Program supported by ONR and managed by Dr. Jack Price (David Taylor Model Basin, NSWCCD). The authors are grateful to Dr. Mike Levine for internal support of this work.

Over the years of the development, many colleagues have influenced and contributed to our work. The authors would like acknowledge Prof. Pol Spanos (Rice University), Dr. Art Reed, Mr. Tim Smith, Mr. Brad Campbell, Dr. Mike Hughes (David Taylor Model Basin, NSWCCD), and Prof. Ross Leadbetter (University of North Carolina Chapel Hill).

7. REFERENCES

Belenky, V.L. 2011, "On Self-Repeating Effect in Reconstruction of Irregular Waves", Chapter 33 of "Contemporary Ideas on Ship Stability", Neves, M.A.S., *et al.* eds., Springer, pp. 589-598.

Belenky, V., Pipiras, V., Kent, C., Hughes, M., Campbell, B. and Smith, T. 2013, "On the Statistical Uncertainties of Time-domain-based Assessment of Stability Failures: Confidence Interval for the Mean and Variance of a Time Series", Proc. 13th Intl. Ship Stability Workshop, Brest, France, pp. 251-258.

Belenky, V., Weems, K. and Lin, W.M., 2015 "Split-time Method for Estimation of Probability of Capsizing Caused by Pure Loss of Stability" in Proc. 12th Intl. Conf. on Stability of Ships *and Ocean Vehicles STAB 2015*, Glasgow, UK

Bickel, J. P. and Doksum, K. A., 2001 Mathematical Statistics: Basic Ideas and Selected Topics. Vol. 1, Prentice-Hall, Upper Saddle River, NJ.

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

Priestley, M. B., 1981, Spectral Analysis and Time Series, Vol. 1, Academic Press, New York.

Reed, A., 2011, "26th ITTC Parametric Roll Benchmark Study", Proc. 12th Intl. Ship Stability Workshop, Washington DC, USA pp 195-204.

Smith, T. C. and Zuzick A. 2015, "Validation of Statistical Extrapolation Methods for Large Motion Prediction" in Proc. 12th Intl. Conf. on Stability of Ships and Ocean Vehicles STAB 2015, Glasgow, UK.