On the Statistical Uncertainties of Time-domain-based Assessment of Stability Failures: Confidence Interval for the Mean and Variance of a Time Series

Vadim Belenky[1], Vladas Pipiras[2], Christopher Kent[1], Michael Hughes[1], Bradley Campbell[1], Timothy Smith[1]

[1] *David Taylor Model Basin – NSWCCD*

[2] *University of North Carolina at Chapel Hill*

ABSTRACT

The paper addresses one of the critical elements of statistical uncertainty of simulated or measured roll motions – confidence interval of the variance estimate. The paper revisits the derivation of the formula for the variance of the sample variance of a stationary stochastic process in order to re-examine the assumptions, especially the one related to the process having a normal distribution. The relation between the formula and the confidence interval based on treating the variance estimate of different records as separate data points is also considered.

KEYWORDS

Roll motions, statistical uncertainty, time-domain simulations

INTRODUCTION

With the development of advanced hydrodynamic codes capable of predicting very nonlinear roll motions, there is an opportunity for the time domain assessment of dynamic stability to become a part of the design process. While addressing the issue of the nonlinearity of large-amplitude motion, time domain simulations create the issue of statistical uncertainty. A time domain simulation of ship motions in irregular seas is a Monte-Carlo method, so any result derived from them (such as the variance of a mode of motion) is a random number. The same is true for the results of model tests in irregular waves and full scale seakeeping trials.

Since the random nature of these results is inherent and cannot be avoided, it is essential to characterize the uncertainty and make it a part of the design analysis. Characterization of the statistical uncertainty of these results is the main objective of the paper.

While evaluation of the confidence interval of the sample variance is one of the most basic statistical problems, there are several details that tend to complicate its evaluation.

First, the inertia of the ship leads to statistical dependence between successive points of the motion time series.

Second, the process of large-amplitude roll response is nonlinear and cannot be assumed to be Gaussian. Unfortunately, the "standard" formulae for the confidence interval use this assumption in one way or another. Parametric roll is a good example of such a process; see *e.g.* Hashimoto *et al* (2006).

Third, the nonlinearity of stability-related problems may lead to the practical inapplicability of the ergodic assumption when multiple records are required to carry out the analysis. This problem became particularly clear while attempting to compare parametric roll results (Reed, 2011).

Thus, the method of characterization of statistical uncertainty of the results of the time-

domain numerical simulations (or model tests in irregular waves) must be able to treat these three features of large-amplitude roll motion: dependence, non-Gaussian distribution and practical non-ergodicity.

In an attempt to account for dependence, Belenky & Weems (2008) used a standard formula (Priestley, 1981) where the estimate of the autocorrelation function was introduced to handle the data dependence. Practical non-ergodicity is addressed by considering several records of roll, as was done in (Reed, 2011). The "last mile" is the assumption of the Gaussian distribution used in standard formula in all the cited works. The focus of this paper is to understand influence of this assumption and see if it can be avoided.

THEORETICAL ANALYSIS

Measure of Uncertainty

The calculation of the confidence interval of a statistical quantity requires an assumption of the distribution of that quantity. With few exceptions (a mean value estimate of the normal variable follows the Student *t*-distribution, while the estimate of the variance has a chi-square distribution), these distributions are unknown.

The assumption that the estimate follows the normal distribution is based on the central limit theorem, since estimates involve summation of random numbers. The caveat is that the sample size should be large enough, as the central limit theorem, strictly speaking, addresses a limiting distribution (as hinted by its name). One should be especially careful applying the normal distribution for the variance estimate as the variance is a positive value by definition, while the normal distribution also supports negative numbers. Nevertheless, if the sample size is large enough, the confidence interval is expected to be relatively small and the influence of asymmetry of the real distribution of the estimate may be neglected. The sample size is expected to be large, because several records are needed to handle practical non-ergodicity.

Once the assumption of the normal distribution of the variance estimate is accepted, the variance of the variance is the only value needed to calculate the confidence interval.

Variance of Mean Value Estimate

Priestley (1981) gives a general direction on the derivation of the formulae for the mean value and variance estimates. This derivation is reproduced here, in order to understand the necessity and role of the Gaussian assumption for the distribution of the process. Consider the variance of the mean value estimate $\hat{m}$ (the symbol above means "estimate") of a stationary process x represented as a record with N points without any further assumptions.

$$Var(\hat{m}) = Var\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} Cov(x_i, x_j) \tag{1}$$

where $Var(..)$ is the variance operator and $Cov(..)$ is the covariance operator. Equation (1) is a standard one; it expresses the variance of a sum of dependent random variables. Since the process x is assumed 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:

$$Cov(x_i, x_j) = R(t_{i-j}) = R(\tau_k) \qquad k = 0, 1..., N-1 \tag{2}$$

Consider a sum of all the elements of the covariance matrix that are needed to compute the variance of the mean estimate in Equation (1):

$$\sum_{i=1}^{N}\sum_{j=1}^{N} Cov(x_i, x_j) = \sum \begin{pmatrix} R(\tau_0) & R(\tau_1) & \dots & R(\tau_{N-2}) & R(\tau_{N-1}) \\ R(\tau_1) & R(\tau_0) & \dots & R(\tau_{N-3}) & R(\tau_{N-2}) \\ \dots & \dots & \dots & \dots & .. \\ R(\tau_{N-2}) & R(\tau_{N-3}) & \dots & R(\tau_0) & R(\tau_1) \\ R(\tau_{N-1}) & R(\tau_{N-2}) & \dots & R(\tau_1) & R(\tau_0) \end{pmatrix} \tag{3}$$

Note that all the elements of the main diagonal of the covariance matrix are the same and equal to variance of the process V, since the auto-covariance function calculated for τ_0=0 is the variance:

$$R(\tau_0) = R(0) = Var(x) = V \qquad (4)$$

In fact all the 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 $N\times N$ square matrix contains N elements; the line of elements parallel to the main diagonal and located next to it, contains only N-1 elements. Each next line will have one element less, until it comes to the low-left or upper-right corner with one element only. Thus the sum in Equation (1) can be presented as (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):

$$\sum_{i=1}^{N}\sum_{j=1}^{N} Cov(x_i, x_j) = N \cdot V + \qquad (5)$$
$$2\left((N-1)R(\tau_1) + (N-2)R(\tau_2) + \ldots + R(\tau_{N-1})\right) =$$
$$N \cdot V + 2 \cdot \sum_{i=1}^{N-1} (N-i) R(\tau_i)$$

Substitution of Equation (5) into Equation (1) leads to the standard formula for the variance of the mean value estimate (see *e.g.* Priestly, 1981)

$$Var(\hat{m}) = \frac{V}{N} + \frac{2}{N} \cdot \sum_{i=1}^{N-1} \left(1 - \frac{i}{N}\right) \cdot R(\tau_i) \qquad (6)$$

The first term in Equation (6) is actually a variance of the mean estimate of the 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 one for the random variable, because the auto-covariance function of the white noise equals zero for all non-zero time lags.

Variance of Variance Estimate

By definition the variance is the average of centered squares, thus a process y is introduced as:

$$y_i = (x_i - m)^2 \approx (x_i - \hat{m})^2 \qquad (7)$$

Then the estimate of the mean value of the process y is the estimate of the variance of the original process x:

$$\hat{m}_y = \hat{V} \qquad (8)$$

Then the variance of the mean estimate of the process y is the variance of the variance estimate of the process x:

$$Var(\hat{V}) = \frac{V_y}{N} + \frac{2}{N} \cdot \sum_{i=1}^{N-1} \left(1 - \frac{i}{N}\right) \cdot R_y(\tau_i) \qquad (9)$$

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

This is the place when the assumption of the Gaussian distribution for the process x is made in order to arrive at the standard formula of the variance of the variance estimate. If the process x has a normal distribution:

$$V_y = 2 \cdot V^2$$
$$R_y(\tau) = 2 \cdot (R(\tau))^2 \qquad (10)$$

Substitution of (10) into (9) leads to the standard formula for variance of the variance estimate (see *e.g.* Priestly, 1981):

$$Var(\hat{V}) = \frac{2V^2}{N} + \frac{4}{N} \sum_{i=1}^{N-1} \left(1 - \frac{i}{N}\right) (R(\tau_i))^2 \qquad (11)$$

Equation (11) can also be expressed in an alternative form where the symmetric properties of the covariance matrix are not used. This form was used, for example, in (Reed, 2011):

$$Var(\hat{V}) = \frac{2}{N} \sum_{i=-(N-1)}^{N-1} \left(1 - \frac{|i|}{N}\right) (R(\tau_{|i|}))^2 \qquad (12)$$

Note that (12) does not have an explicit term that includes the variance, but since the index of the time lag goes through zero, this term is, indeed, included.

It seems that there is no apparent reason to use the Gaussian assumption. The calculation of

the auto-covariance function of the centered squares requires a little additional computation effort in comparison with straight auto-covariance function.

Variance of Ensemble Variance

Consider an ensemble of *Nr* records, each with N_i data points. The time increment is assumed to be the same for all the records, which is the usual practice for both numerical simulations and model tests.

Then the statistical weight for each record is expressed as follows

$$W_i = \frac{N_i}{\sum_{i=1}^{Nr} N_i} = \frac{N_i}{N_{total}} \tag{13}$$

where N_{total} is the total number of points in the ensemble. The ensemble estimate for the mean value is calculated for all the points

$$\hat{m}_a = \frac{1}{N_{total}} \sum_{i=1}^{Nr} \sum_{j=1}^{N_i} x_{i,j} = \frac{1}{N_{total}} \sum_{i=1}^{Nr} \frac{N_i}{N_i} \sum_{j=1}^{N_i} x_{i,j} = \sum_{i=1}^{Nr} W_i \left(\frac{1}{N_i} \sum_{j=1}^{N_i} x_{i,j} \right) = \sum_{i=1}^{Nr} W_i \hat{m}_i \tag{14}$$

where $\hat{m}_i$ is the mean value estimate for a record *i*. The data point $x_{i,j}$ in Equation (14) has two indexes *i* for the record and *j* is the index within a record. Since the records can be of different lengths, the set of data points $x_{i,j}$ do not constitute a matrix. The ensemble estimate for the variance is expressed analogously to the mean value:

$$\hat{V}_a = \sum_{i=1}^{Nr} W_i \hat{V}_i \tag{15}$$

where $\hat{V}_i$ is the variance estimate for record *i*. The variance of the ensemble variance estimate can be calculated as:

$$Var\left(\hat{V}_a\right) = \sum_{i=1}^{Nr} W_i^2 Var(\hat{V}_i) \tag{16}$$

where the variance of the variance estimate for each record is taken from Equation (9).

Direct Estimate of Variance of the Variance

Consider the variance estimate of each record as a realization of a random number. The average variance of the record estimate is (accounting for the fact that each variance estimate with the ensemble has a statistical weight W_i,)

$$\hat{Var}\left(\hat{V}\right) = \sum_{i=1}^{Nr} W_i \left(\hat{V}_i - \hat{V}_a\right)^2 \tag{17}$$

Equation (17) is not equivalent to Equation (16); it gives the average variance of each record, so it should be equivalent to Equation (9) averaged through the ensemble. The variance of the ensemble estimate should be treated as the variance of the mean of the record estimates:

$$\hat{Var}\left(\hat{V}_a\right) = \sum_{i=1}^{Nr} W_i^2 \left(\hat{V}_i - \hat{V}_a\right)^2 \tag{18}$$

Substituting Equation (8) into (18):

$$\hat{Var}\left(\hat{V}\right) = \sum_{i=1}^{Nr} W_i \left(\hat{m}_{yi} - \hat{V}_a\right)^2 = \sum_{i=1}^{Nr} W_i \left(\frac{1}{N_i} \sum_{j=1}^{N_i} y_{i,j} - \hat{V}_a \right)^2 = \sum_{i=1}^{Nr} \frac{W_i}{N_i^2} \left(\sum_{j=1}^{N_i} \left(y_{i,j} - \hat{V}_a \right) \right)^2 \tag{19}$$

Using the known formula for the square of a sum, one can write:

$$\hat{Var}\left(\hat{V}\right) = \sum_{i=1}^{Nr} \frac{W_i}{N_i} \left(\frac{1}{N_i} \sum_{j=1}^{N_i} \left(y_{i,j} - \hat{V}_a \right)^2 + \frac{2}{N_i} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i - j} \left(y_{i,j} - \hat{V}_a \right) \left(y_{i,j+k} - \hat{V}_a \right) \right) \tag{20}$$

The second term in equation (20) can be considered as an estimate for the auto-covariance function of a single record of centered squares that uses population mean (15) instead of a record mean. It has to be distinguished from the estimate based on the record data only:

$$\hat{R}_{py}\left(\tau_j\right) = \frac{1}{N_i - j} \sum_{k=1}^{N_i - j} (y_{i,j} - \hat{V}_a)(y_{i,j+k} - \hat{V}_a) \tag{21}$$

The first term in the formula (20) is the same estimate auto-covariance at zero time lag. Thus

$$Var\left(\hat{V}\right)= \sum_{i=1}^{Nr} W_i \left(\frac{\hat{R}_{py}(0)}{N_i} + \frac{2}{N_i} \sum_{j=1}^{N_i} \left(1 - \frac{j}{N_i}\right) \hat{R}_{py}\left(\tau_j\right) \right) \qquad (22)$$

Equation (22) is similar Equation (9). The difference is that not only averaging over all the records in the ensemble, but also uses the population mean instead of the record mean for calculation of the auto-covariance function. Thus direct estimate of the variance of variance (17) is equivalent to population average of the record variance of variance, where population mean is used for evaluation of auto-covariance function of the centered squares.

NUMERICAL ANALYSIS

Source of Ship Roll Data

A hybrid model (Weems & Wundrow, 2013) was used to reproduce roll motion as a fast and easy way to reproduce roll motions with the correct type of nonlinearity. The model calculates the Froude-Krylov and hydrostatic forces on the actual submerged volume for three degrees of freedom: heave, roll and pitch. Calculations were performed for the ONR tumblehome topside configuration (Bishop, *et al*, 2005); this configuration is representative of an unconventional hull design and produces sufficiently nonlinear motions bringing into question the Gaussian assumption for roll motions while assessing statistical uncertainty.

The motions were simulated for a sea state described by a significant wave height of 7.5 m and a modal period of 15s. Long-crested irregular waves were modeled with the Bretschneider spectrum. The speed was 6 knots in stern-quartering seas (45 degrees). The spectrum was discretized with 100 uniformly distributed frequencies that facilitated modeling 10 minute long records. The ensemble (population) consisted of 300 records totaling 50 hrs worth of data.

Estimation of Auto-Covariance

Strictly speaking, only the auto-covariance function for centered squares is needed for Equation (9), however, it may be instructive to look at the auto-covariance of the original process as well. The formal definition of the auto-covariance estimate is given in Equation (21) and rewritten here for the process x

$$\hat{R}(\tau_i) = \frac{1}{N-i} \sum_{j=1}^{N-i} (x_j - \hat{m})(x_{j+i} - \hat{m}) \qquad (23)$$

When the time lag τ_I becomes large, the volume of the sample available for averaging decreases dramatically. From Figure 1, an increase in the magnitude of the auto-covariance function for the large time lags can be observed.

This loss of accuracy can be alleviated by a simple weighting factor: $(N\text{-}i)/N$, re-writing Equation (23) as follows:

$$\hat{R}(\tau_i) = \frac{1}{N} \sum_{j=1}^{N-i} (x_j - \hat{m})(x_{j+i} - \hat{m}) \qquad (24)$$

The weighting results in little change to the auto-covariance function for small time lags as the difference between N and $N\text{-}i$ 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 result of weighting the estimate of the auto-covariance function is shown in Fig 2.

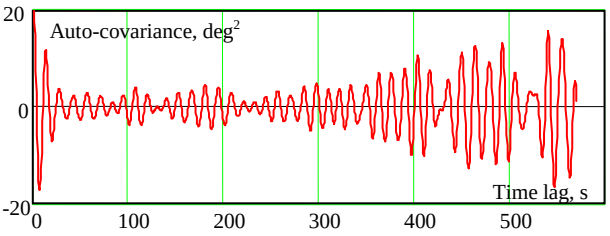

Fig. 1: Auto-covariance function estimated from a single record using Equation (23).

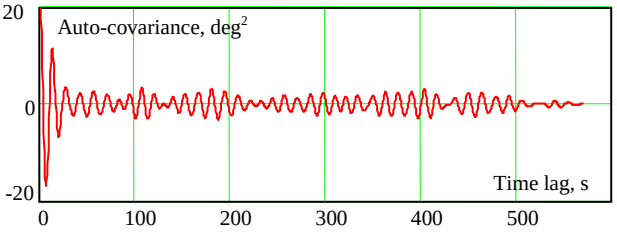

Fig. 2: Auto-covariance function estimated from a single record using Equation (24) using linear weighting factor.

Comparing Figures 1 and 2, one can see that the initial part did not change much, however the amount of numerical "noise" has decreased significantly. Averaging the estimate across the records further decreases this noise and accounts for possible practical non-ergodicity:

$$\hat{R}_a(\tau_i) = \sum_{i=1}^{Nr} W_i \hat{R}(\tau_i) \qquad (25)$$

where Nr is the total number of records, N_i the number of points in i-th record, W_i is a weighing factor of i-th record. Figure 3 shows the estimate of the auto-covariance function averaged for 300 records. As expected, the "noise" is practically gone.

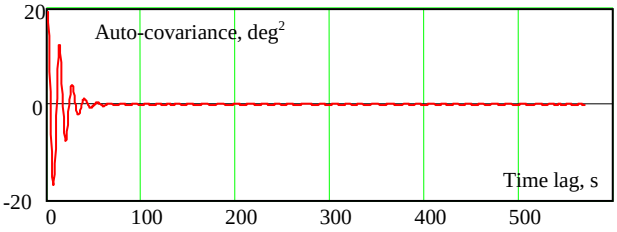

Fig. 3: Averaged auto-covariance function, Equation (15).

Estimation of Auto-Covariance for the Centered Squares

The estimation of the auto-covariance function for the centered squares process is similar; first the weighted record estimate is calculated, then the population average is evaluated.

$$\hat{R}_y(\tau_i) = \frac{1}{N}\sum_{j=1}^{N-i}(y_j - \hat{V})(y_{j+i} - \hat{V})$$
$$\hat{R}_{ay}(\tau_i) = \sum_{i=1}^{Nr} W_i \hat{R}_y(\tau_i) \qquad (26)$$

Figure 4 shows the population average for the auto-covariance of the centered squares, while Figure 5 contains the zoomed-in view of the first 200 seconds of the estimate.

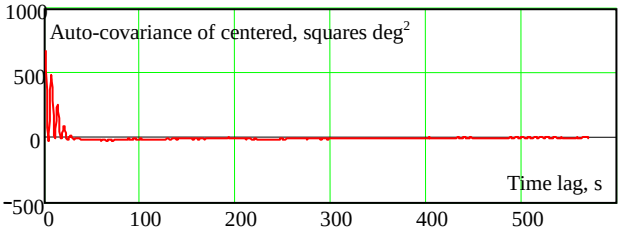

Fig. 4: Averaged auto-covariance function of the centered squares, Equation (27)

The shape of the auto-covariance function of the centered squares is drastically different compared with the auto-covariance function of the original process. The folds are located mostly on the positive side and there is a negative "tail" slowly approaching zero. The appearance of the negative "tail" is not a result of numerical error, but a consequence of mostly positive folds; it comes from the known property of the auto-covariance function:

$$R(\tau_0) + 2\sum_{i=1}^{N-1} R(\tau_i) = 0 \qquad (27)$$

If the folds are mostly positive, the rest of the auto-covariance must be negative to bring the sum (27) to zero.

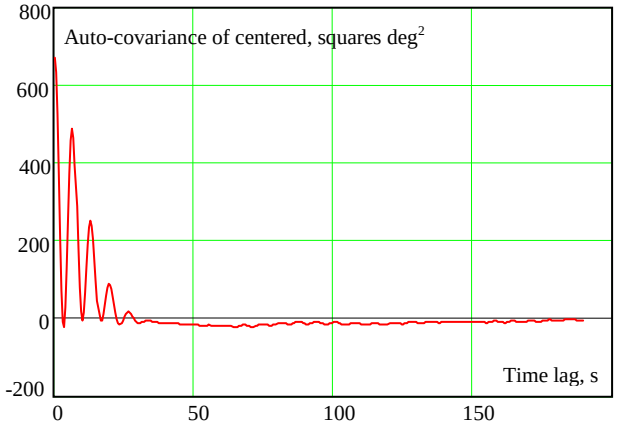

Fig. 5: Averaged auto-covariance function of the centered squares, zoomed-in view

Possible Scheme of Calculation of the Variance of Variance Estimate

Since the large-amplitude roll response may be practically non-ergodic, it makes sense to use the ensemble/population estimate whenever possible. Thus the mean value estimate (15) should be calculated first to be used for further estimates.

Then the centered squares are calculated:

$$y_{i,j} = (x_{i,j} - \hat{m}_a)^2 \qquad (28)$$

The mean value estimate for centered squares is the variance estimate for the original process

$$\hat{V} = \hat{m}_y = \sum_{i=1}^{Nr} W_i \left(\frac{1}{N_i} \sum_{j=1}^{N_i} y_{i,j} \right) \qquad (29)$$

The auto-covariance of the centered squares is calculated with Equation (21) and averaged over the population:

$$\hat{R}_{apy}(\tau_i) = \sum_{i=1}^{Nr} W_i \hat{R}_{py}(\tau_i) \qquad (30)$$

Then the variance of the variance estimate for each record needs to be calculated. To decrease variability for larger time lags, it is proposed to remove the summands above the cutoff point *M*:

$$\hat{Var}(\hat{V}_i) = \frac{\hat{R}_{apy}(0)}{N_i} + \frac{2}{N_i}\sum_{j=1}^{M-1}\left(1-\frac{j}{M}\right)\hat{R}_{apy}(\tau_j) \qquad (31)$$

It is proposed to set the cutoff point *M* to half of the average number of points of a record. In any case, for the correct definition of the ensemble- averaged auto-covariance function of centered squares (26):

$$M \le \min(N_i) \qquad (32)$$

The final result is the variance of the variance estimate for the ensemble that is calculated with Equation (16). Once the variance of the variance estimate is calculated, the last step is the assessment of the confidence interval. Since the estimate is assumed to be distributed normally, the half-width of the confidence interval as expressed as:

$$\Delta\hat{V}_a = \varepsilon_\beta \sqrt{\hat{Var}(\hat{V}_a)} \qquad (33)$$

Where ε_β is a coefficient dependent on the accepted confidence probability β, *e.g*:

$$\beta = 0.95 \quad ; \quad \varepsilon_\beta = 1.96$$

Figure 6 shows a comparison of three different ways to compute the ensemble/ population estimate of the variance with confidence interval.

The "standard" Gaussian assumption overestimates uncertainty compared to the two other methods. Equation (31) also shows a slight overestimation compare to the direct estimate (18). However, more calculations are needed to conclude that the observed differences are of general nature.

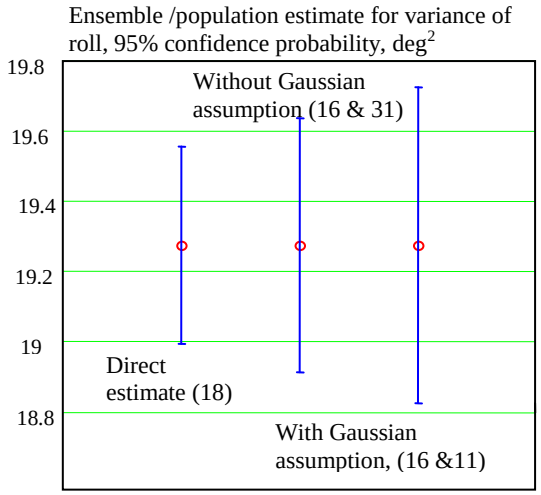

Fig. 6: Comparison of different methods to compute confidence interval on the ensemble variance estimate. Equations in parentheses.

CONCLUSIONS AND FUTURE WORK

Contrary to popular opinion, the derivation of the formula for variance of the variance estimate is not bulky and is quite straight-forward. The assumption of the Gaussian distribution of the process is actually not necessary, if one can estimate a covariance function of centered squares of the process.

Direct estimation when the variance of each record is considered as a separate data point is similar to the formula of variance of the variance. The difference includes use of the population mean instead of the record mean for the centered squares. Applying a linear weighting function on the estimate of the auto-covariance function helps to significantly reduce statistical "noise" caused by the decrease of available data in large time lags.

The next logical step is to test these calculations. This would include creating a large set of ensembles in order to see how well the computed confidence interval captures the expected number of ensemble estimates. The fraction of estimates falling with the confidence interval should be close to the given confidence probability.

ACKNOWLEGEMENTS

The work described in this paper has been funded by the Office of Naval Research under Dr. Patrick Purtell and Dr. Ki-Han Kim. This support is gratefully acknowledged by the authors. Participation of Dr. Pipiras was supported by the summer faculty program supported by ONR and managed by Dr. Jack Price of David Taylor Model Basin.

The authors are grateful to Dr. D. Drazen and Dr M. Levine for internal support of this work. Discussions with Dr. A. Reed have been very fruitful.

REFERENCES

Belenky, V. & K.M. Weems (2008) Procedure for Probabilistic Evaluation of Large Amplitude Roll Motions in *Proc. of the Osaka Colloquium on Seakeeping and Stability of Ships*, Osaka, Japan

Bishop, R. C., W. Belknap, C. Turner, B. Simon, J. H. Kim (2005) *Parametric Investigation on the Influence of GM, Roll damping, and Above-Water Form on the Roll Response of Model 5613*. Report NSWCCD-50-TR-2005/027.

Hashimoto, H., N. Umeda, A. Matsuda, & S. Nakamura, (2006), "Experimental and Numerical Studies on Parametric Roll of a Post-Panamax Container Ship in Irregular Waves," *Proc. 9th Intl. Conf. on the Stability of Ships and Ocean Vehicles*, Rio de Janeiro, Brazil, Vol. 1, p 181-190.

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.

Weems, K., & D. Wundrow (2013) "Hybrid models for fast time-domain simulation of stability failures in irregular waves with volume-based calculations for Froude-Krylov and Hydrostatic Forces"; *Proc. 13th Intl. Ship Stability Workshop*, Brest, France