POAC'13
Espoo, Finland

Proceedings of the 22nd International Conference on
Port and Ocean Engineering under Arctic Conditions
June 9-13, 2013
Espoo, Finland

STATISTICALLY BASED METHOD OF REPRESENTING ICE LOADS SIGNAL AS A SUM OF SEVERAL UNCORRELATED AND STATIONARY PROCESSES

Petr N. Zvyagin[1,2], Kirill E. Sazonov[1]
1. Krylov State Research Centre, St. Petersburg, RUSSIA
2. St. Petersburg State Polytechnical University, St. Petersburg, RUSSIA

ABSTRACT

In the paper the method of signal processing is offered. It was developed specially for processing of ice loads observations made in ice basin. The proposed method includes numerical search of the signal oscillatory components parameters by minimization of these components covariance and model discrepancy variance. In the method the clustering procedure is used for the final frequencies estimation.

Uncorrelated random processes which parameters are estimated this way appear to be useful to describe the greater part of initial signal variance. The components of particular interest are stationary stochastic processes with determined frequency and random phase. The discrepancy for such model is in general a non-stationary process, but with variance much less then variance of initial signal. The characteristic of this method is that it uses not large time intervals, for example 1 second. The selection of time interval length depends on existing noise characteristics.

In the paper the offered method is applied to explain the ice loads measurements that were got in experiments with tanker and ice-breaker hulls in ice basin. In the last section of the paper some statistics values for estimated parameters and ice load signal discrepancy are given.

Comparing to spectral investigation, this method do not use necessarily orthogonal trigonometric functions to describe oscillatory components. But at the same time numerically found frequencies conform to the general spectral investigation results.

STOCHASTIC MODEL FOR ICE LOADS SIGNAL

Consider the dynamometer signal $X(t)$ that was received in the experiment with the model of tanker or ice-breaker hull in the ice basin. Assume that the values of the signal recorded in an equidistant discrete time points, so we can take $X(t)$ as a stochastic process with discrete time and continuous state, i.e. each random variable $X(t_0)$ for the time value t_0 is continuous.

The recorded signal is a superposition of model response to the phenomenon of ice destruction, equipment noise and the model's swinging and pitching due to the forces applied on its hull (Sazonov, 2010). Some of these components on a relatively short period of time can be described by a model of stationary signal.

Let fix the time window of n discrete observations starting from the moment of time t_k. Let this time window corresponds to T seconds. Let we select the time window that way that within it the process $X(t)$ can be represented by the following expansion:

$$X_k(t) = m_k + \sum_{j=1}^{q} V_j(t) + \sum_{i=1}^{p} U_i(t) + Y(t), \qquad t = t_k, t_{k+1}, \ldots, t_{k+n} \ . \qquad (1)$$

Here $m_k = const$; $V_j(t)$, $j = 1..q$, $U_i(t)$, $i = 1..p$, and $Y(t)$ - uncorrelated stochastic processes, and the process $\sum_{j=1}^{q} V_j(t) + \sum_{i=1}^{p} U_i(t) + Y(t)$ is centered. Assume

$$V_j(t) = A_j \cdot \sin(\omega_j t + \varphi_j(t)), \qquad (2)$$

$$U_i(t) = B_i \cdot \sin(\omega_i t + \Phi_i), \qquad (3)$$

where A_j, B_i and Φ_i – random variables, Φ_i has a uniform distribution on the interval $[0;2\pi]$; ω_i – constant parameters, $\varphi_j(t)$ – nonrandom functions. Here processes $U_i(t)$ are stationary in wide sense. Concerning the processes $V_j(t)$, in general they can not be considered as stationary.

Processes $V_j(t)$, $j = 1..q$, $U_i(t)$, $i = 1..p$, and $Y_1(t)$ are uncorrelated by assumption, hence

$$DX(t) = \sum_{j-1}^{q} DV_j(t) + \sum_{i=1}^{p} DU_i(t) + DY(t), \qquad (4)$$

where $DX(t)$ is variance of $X(t)$.

We suggest the following method of estimating the parameters of (1).

Suppose we have a realization of the random process $X(t)$ at successive times $t = 0,1,\ldots,N$, $N >> n$. Consider the part of this realization which bounded by a the time window starting at the point of time $t_k = k$. Then $\bar{x}_k = \frac{1}{n}\sum_{i=0}^{n-1} X(t_{k+i})$ is consistent, unbiased and effective estimator of m_k.

Assume the validity of expression (1). Denote $\tilde{X}(t) = X(t) - \bar{x}_k$. For the realization $\tilde{X}(t)$ bounded by the time window with index k, that starts at moment of time t_k, we can numerically solve the next optimization problem:

$$s^2(E) \underset{b,\omega,\varphi}{\to} \min, \qquad \hat{K}(U(b,\omega,\varphi), E) \underset{b,\omega,\varphi}{\to} \min \qquad (5)$$

$$E_k(b,\omega,\varphi) = \sum_{i=0}^{n-1} \left(\tilde{X}(t_{k+i}) - U(b,\omega,\varphi)\right),$$

$$U(b,\omega,\varphi) = b \cdot \sin(\omega t + \varphi), \qquad (6)$$

with boundary $b \geq 0$, where ω, b, φ are real variables. Here m^2 is a sample variance and $\hat{K}$ is two-sample covariance. For this problem we take $\omega \in [0,\pi]$, $\varphi \in [0;2\pi]$, $b \in \mathrm{R}$.

Denote *m* the number of time windows of the same length that are applied to realization of $X(t)$, thus $k = 1..m$. Solving the problem (5) for *m* time windows we obtain samples $\{b\}$, $\{\omega\}$, $\{\varphi\}$ of size *m* as well as segments $E_k(t)$, $k = 1..m$ of realizations of error process.

We note that there is no reason to consider $E_k(t)$ as realization of stationary process as well as it is for the original signal $X(t)$. But from $E_k(t)$ in some cases, we can allocate stationary component with significant variance.

By the statement of the problem (5) the components E and U whose parameters have been estimated numerically typically have low covariance. Therefore the use of the theoretical formula (4) in most practical cases is justified.

We solve the problem (5) *r* times. In all cases except the first, instead of $X(t)$ we'll take the corresponding error after removal out of the original signal the components that were previously evaluated by (6).

We will unite samples $\{\omega\}_j$, $j = 1..r$, in a total sample $\{\Omega\}$. For the representativeness of the sample $\{\Omega\}$ we will take *m* and *r* so that their product mr will be not less than 50. Here mr is the size of the sample $\{\Omega\}$.

There is reason to assume the presence of narrow-band harmonic components for a stochastic process of ice loads, recorded when dragging a model of a tanker or ice-breaker through an ice field. We will implement clustering on $\{\Omega\}$ by a given number of clusters. For this purpose we will use K-means clustering algorithm also known as Hard C-means (HCM) (McKay, 2003 & Rutkowski, 2008).

As estimators of parameter ω_i of components (3)-(4) we'll take centers of the clusters that have been found by the mentioned algorithm.

If the elements of the cluster considered as a separate observations of the random variable W, then, for example, the mean ω^* of the elements of this cluster will have properties of consistent, unbiased and efficient estimator of the W expectation $M(W)$. We will use K-means modification that takes arithmetic mean of elements of the cluster as its center.

Numerical experiments have shown that the composition of the sample $\{\Omega\}$, in general, depends on the choice of the time window size *n* because of nonstationarity of the original signal $X(t)$. However, for a special type of experiments in the ice basin, there is usually an interval of *n*, that the centers of the clusters are almost constant. However, for the special type of experiment in ice basin usually it is an interval for *n* exist, so that within this interval cluster centers for $\{\Omega\}$ are nearly constant. Evident that this happens because of the following two reasons:

1) if the used time window has a length of *n*, close to the period of existing in $X(t)$ low-frequency oscillations with a variable phase and slightly varying frequency;
2) signal $X(t)$ is noised by the component with constant frequency but variable phase and amplitude.

If we will greatly increase the size of the time window, the frequencies obtained during the solution of problem (5) become very low. This shows that for a very wide window, the ordinary linear regression becomes the preferred trend comparing to the trigonometric function (6).

Then in the problem (5) we shall take ω_i^* as a fixed value of parameter ω_i and after that recalculate values for b_k и φ_k for all time windows, $k = 1..m$. Thus we shall get new samples $\{\tilde{b}\}_i$ and $\{\tilde{\varphi}\}_i$.

Investigating of ice loads that were registered in tests in the ice basin, demonstrated that oscillatory components with constant frequencies can be attributed to one of two types − (2) or (3).

Using the assumption (2) we can consider that values φ_k found for fixed ω_i are depend on k. So the researcher needs to determine the type of function $\varphi(k)$. Experiments have shown that for the "core" frequency ω_i, apparently responsible for the vessel hull's rocking due to ice loads, $\varphi(k)$ sometimes had the character of the oscillatory function.

If the dependence φ_k on k is not evident we can unite φ_k for fixed ω_i, $k = 1..m$, into the sample $\{\tilde{\varphi}\}_i$ of size *m*. According to the assumption (3) we can consider $\{\tilde{\varphi}\}_i$ as a sample generated by random variable Φ_i. Let investigate he sample $\{\tilde{\varphi}\}_i$. For the mentioned problem of ice loads data analysis we will propose a hypothesis - "a random variable Φ_i has the uniform distribution on the interval $[0;2\pi]$". For testing of this hypothesis we can use Kolmogorov-Smirnov test (Bolshev & Smirnov, 1983).

Model (1) has next useful features. If we can ignore components $U_i(t)$, $i = 1..p$, then $m_k + \sum_{j=1}^{q} MA_j \cdot \sin(\omega_j t + \varphi_j(t)) + \sum_{j=1}^{q} \overset{0}{A}_j \cdot \sin(\omega_j t + \varphi_j(t))$, is a so-called canonical decomposition (Pugachev, 1965) of stochastic process $X_k(t) - Y(t)$; here $\overset{0}{A}_j$ are centered random variables A_j. And hence for variance of this process we have:

$$D(X_k(t) - Y(t)) = \sum_{j=1}^{q} DA_j \cdot \sin^2(\omega_j t + \varphi_j(t)).$$

Let consider the component (3). For the expected value of random variable $\sin(\omega t + \Phi_i)$, where Φ_i is uniformly distributed on the interval $[0;2\pi]$, we have:

$$M \sin(\omega t + \Phi_i) = 0 .$$

And therefore in the case of independent random variables B_i и Φ_i, for $U_i(t)$ (3) expected value we have:

$$MU_i(t) = MB_i \cdot M \sin(\omega_i t + \Phi_i) = 0 \qquad (7)$$

Thus in the mentioned case of considering model (1) only with components (3), the error process $Y(t)$ has zero mean, i.e. it is centered.

We can estimate the practical usefulness of offered method using the estimator of the next theoretical ratio:

$$\frac{DY(t)}{DX(t)},$$

In this paper an estimator ρ will be used for this purpose:

$$\rho = \frac{1}{m} \sum_{k=1}^{m} \frac{s^2(Y_k)}{s^2(X_k)} \qquad (8)$$

where $s^2(X_k)$ is a sample variance calculated for the sample of all process $X(t)$ observations within the time window started at t_k; and $s^2(Y_k)$ is sample variance calculated for the sample of errors after deleting all components (2) and (3) within the same time window.

Thus, if the model (2)-(3) is suitable to describe initial signal in the sense of error's variance minimization, then we receive small value of ρ.

At last we can note that the value of $\bar{x}_k$ can be obtained by the regular procedure of smoothing the signal − *m* data items averaging . The variance of $\bar{x}_k, k = 1..m$, will decrease with increasing *m*, i.e. with expansion of time window. But at the same time, because of unstationarity of $X(t)$, we can expect increasing of ρ values for the fixed number *p* of components (2)-(3), and that is undesirable.

APPLICATION OF STOCHASTIC MODEL TO EXPERIMENTAL DATA

We will use the proposed model for the signals that are noised by electrical noise (Figure 1a)) as well as for the signals that are free from such interference (Figure 1b). Let consider experiments with tanker hull model that were conducted in Krylov State Research Center's ice basin. In these experiments model was towed by a trolley with velocity 0,1 m/sec (Figure 1a) and 0,3 m/sec through the ice field of 25 and 40 mm thickness.

Let examine observations of stochastic process $X(t)$ at successive moments of time $t = 0,1,...,1090$. We shall take $n = 90$ as a size of time window, this corresponds to 0,9 seconds.

The time window we shall sequentially shift forward by 10 observations , which corresponds to 0,1 seconds, until the end of $X(t)$, thus $m = 100$.

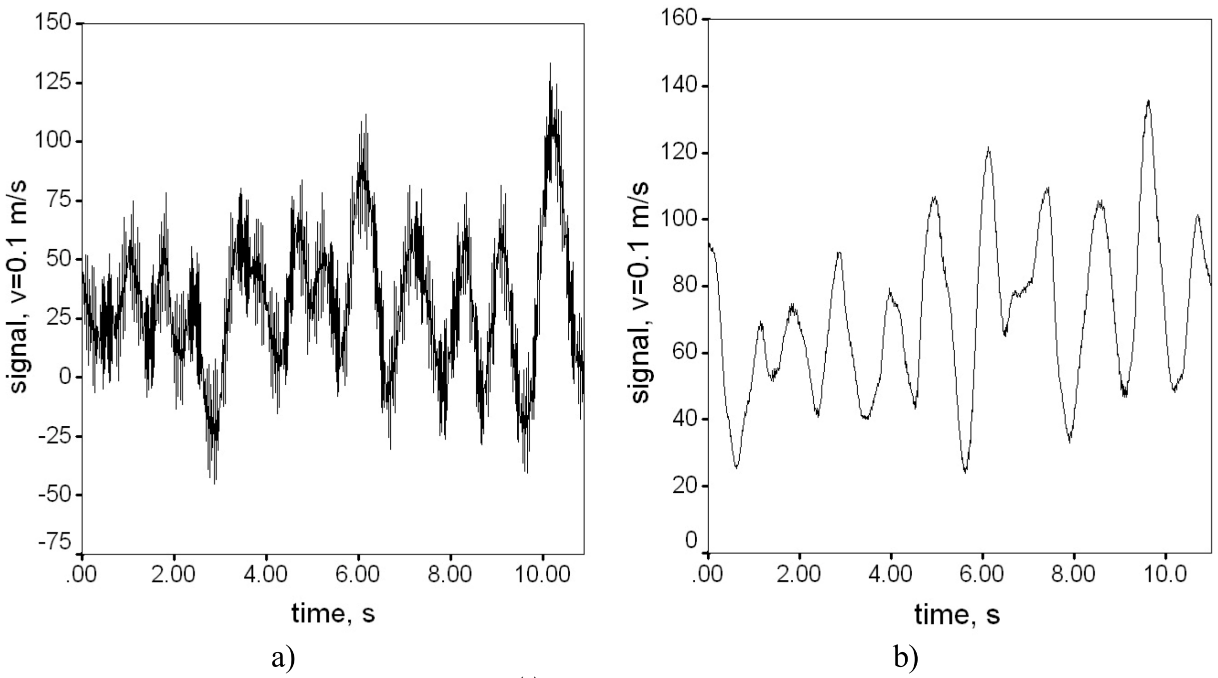


Figure 1. Observations of process $X(t)$, model velocity 0,1 m/sec: a) tanker, ice field thickness is 25 mm (signal is noised by electrical noise); b) ice-breaker, ice field thickness is 50 mm

For the model (1) we shall suppose presence of three components (2) or (3) and therefore we shall solve problem (5) three times for each time window. Frequencies clusters centers have been found by K-means algorithm and they are presented in the Table 1. These results conform to the results of general spectral investigation of time series (Ledermann, 1984). For example, on the Figure 2 the periodogram for the time series from Figure 1a is presented. We can see that the result from the first string of Table 1 obtained just for this time series conforms to the periodogram on Figure 2.

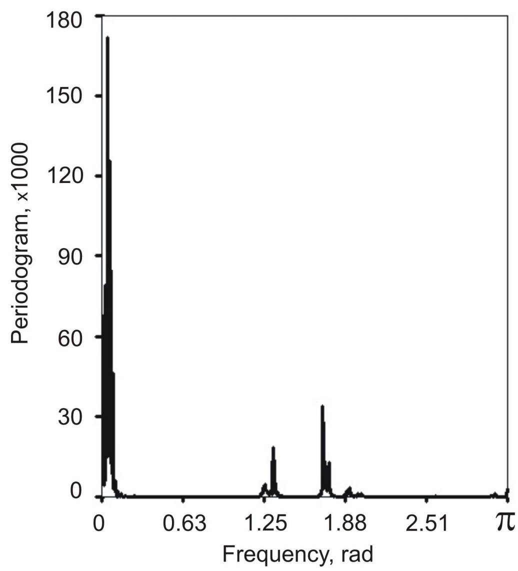


Figure 2. Periodogram for time series presented on Figure 1a.

Table 1. Frequencies for the expression (1), for the model of tanker

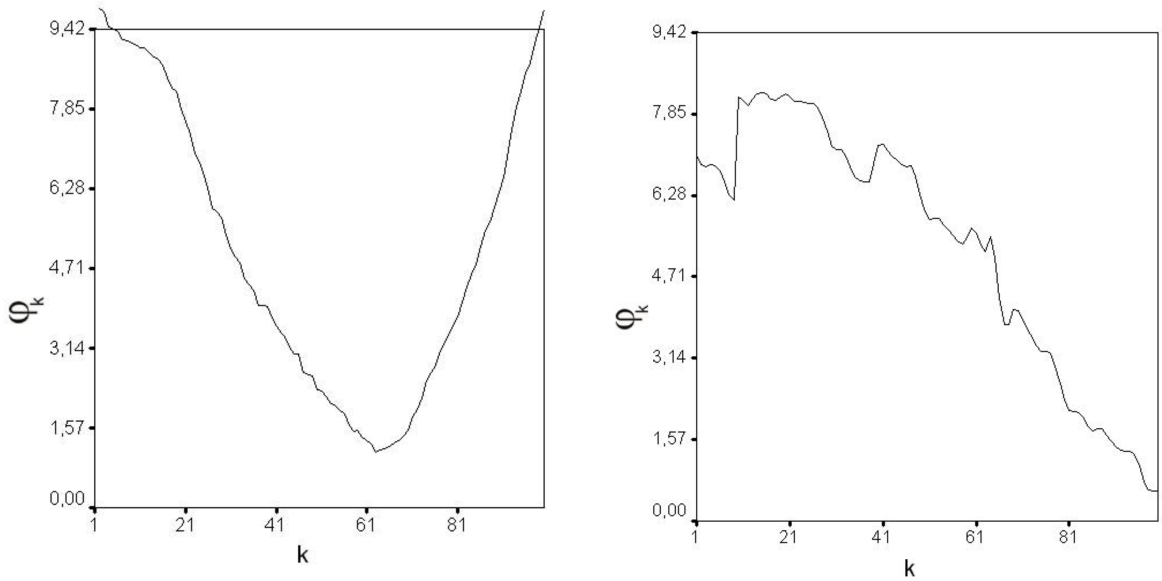
Ice field thickness, velocity of model	Center of the cluster 1 ω_1, radians	Center of the cluster 2 ω_2, radians	Center of the cluster 3 ω_3, radians
25 mm, 0,1 m/sec	0.0685	1.307	1.729
25 mm, 0,3 m/sec	0.0778	1.204	3.033
40 mm, 0,1 m/sec	0.0677	0.37	1.1008
40 mm, 0,3 m/sec	0.0632	0.98	1.4694

Values of the statistic (8) for these 3-component expansions (1) with fixed ω_1, ω_2, ω_3 from Table 1, are given in Table 2. Here ρ characterizes the percentage of unexplained part of the initial signal variance.

Table 2. Statistics ρ (8) for 3-component expansions (1)

Ice field thickness, velocity of tanker's hull model	25 mm, 0,1 m/sec	25 mm, 0,3 m/sec	40 mm, 0,1 m/sec	40 mm, 0,3 m/sec
ρ	0.23	0.32	0.37	0.255

For the electrical interference that noised ice loads signal (Figure 1a), the model (2) was preferred. The plot of $\varphi(k)$ function from (2) for electrical noise is presented on Figure 3a). Values of $\varphi(k)$ depends on time window starting moment, where $t_{starting} = 10k$.



a) b)

Figure 3. Curves for $\varphi(k)$ (phase) in oscillatory component $V_j(t)$ for: a) electricity noise with $\omega = 1{,}729$ extracted from signal on Figure 1a; b) main "pitching" component with frequency $\omega = 0{,}0655$ extracted from signal on Figure 1b,

Consider experiments with icebreaker's hull model. In these experiments the model was towed by a trolley with velocity 0,1 m/sec (Figure 1b) through the ice field of 50 mm thickness. The observed signal was without noise of electrical origin. Centers ω_i, $i = 1,2,3$, for frequency's clusters as well as values of statistics ρ (8) are presented in the Table 3.

Table 3. Frequencies for the expression (1), for the model of ice-breaker

Ice field thickness, velocity model	Center of the cluster 1 ω_1, radians	Center of the cluster 2 ω_2, radians	Center of the cluster 3 ω_3, radians	ρ
50 mm, 0,1 m/sec	0.0655	0.1301	0.222	0.031

Consider centers of clusters that are presented in the table 3. It turned out that for oscillatory component with frequency ω_1 the process (2) corresponds, and for components with frequencies ω_2 and ω_3 the stationary process (3) corresponds. The plot of $\varphi(k)$ function for component with frequency ω_1 is presented on the Figure 3b. This component corresponds to the "core" of the process and supposedly responsible for vessel hull's rocking due to the ice loads (Figure 1b).

Consider hypothesis H_{0i} "the random variable Φ_i is uniformly distributed on the interval $[0;2\pi]$", $i = 2,3$. The sample $\{\tilde{\varphi}\}_i$ values of Kolmogorov-Smirnov statistics λ are presented in the table 4. For the level of significance $\alpha = 0{,}05$ the critical value is $\lambda_{crit} = 1.36$. According to the values presented in table 4 we can accept H_{0i}, $i = 2,3$. So we can consider components with frequencies ω_2 and ω_3 as stationary processes.

Table 4. Results for K-S test of phase Φ_i of $U_i(t)$ stationary component

Ice field thickness, velocity of ice-breaker's hull model	λ for Φ_2	λ for Φ_3
50 mm, 0,1 m/sec	0.4	0.67

OUTCOMES

Thus, we demonstrated the suitability of stochastic model, which is proposed in the article, for description of the processes that can be recorded during the tests of tankers or icebreakers models in the ice basin. The sum of uncorrelated processes with a fixed frequencies can describe a significant part of the initial signal's variance within a time window of about 1 second.

It is possible to suggest that for experiment with towing the vessel's model fixed to the trolley at one point we can attribute main oscillatory processes of the dynamometer signal to the pitching of the model due to the applied ice loads. At the same time the variance of the probable response to a non-stationary ice destruction process was extremely small in a series of experiments with these types of models and method of their fixation.

REFERENCES

1. Bolshev L.N., Smirnov N.V., 1983. Tables of Mathematical Statistics. Moscow: Science. (in Russian), 416 p.
2. Ledermann W. (Editor), 1984. Handbook of Applicable Mathematics: Statistics, Vol. 6, Part B. Wiley, John & Sons, 522 pp.
3. MacKay D., 2003. Information Theory, Inference and Learning Algorithms. Cambridge University Press, 640 pp.
4. Pugachev V.S., 1965. Theory of Random Functions and its Application to Control Problems, Pergamon (Translated from Russian), 833 pp.
5. Rutkowski L, 2008. Computational Intelligence. Methods and Techniques. Springer-Verlag, 514 pp.
6. Sazonov K.E., 2010. The Theoretical Foundations of Ship Operations in Ice. St. Petersburg: Krylov State Research Center Publishing (in Russian), 274 pp.