# Likelihood updating of random process load and resistance parameters by monitoring

Peter Friis-Hansen & Ove Ditlevsen
*Section of Maritime Engineering, Department of Mechanical Engineering,*
*Technical University of Denmark, DK 2800 Lyngby, Denmark*

Keywords: Parametric spectral estimation, likelihood of spectral parameters, response monitoring

ABSTRACT: Spectral parameters for a stationary Gaussian process are most often estimated by Fourier transformation of a realization followed by some smoothing procedure. This smoothing is often a weighted least square fitting of some prespecified parametric form of the spectrum. In this paper it is shown that maximum likelihood estimation is a rational alternative to an arbitrary weighting for least square fitting. The derived likelihood function gets singularities if the spectrum is prescribed with zero values at some frequencies. This is often the case for models of technically relevant processes. The numerical problem caused by these singularities is easily overcome by adding simulated low intensity white noise to the realization. Without changing its parameters the spectrum is hereby lifted above zero by an amount equal to the white noise intensity. The knowledge of an explicit likelihood function, even though it is of complicated mathematical form, allows an approximate Bayesian updating and control of the time development of the parameters. Some of these parameters can be structural parameters that by too much change reveal progressing damage or other malfunctioning. Thus current process monitoring and updating, for example administered in a Bayesian network system, can be a useful aid for the operation of a complicated technical system (large important structure, ship, wind power engine, etc.).

## 1 INTRODUCTION

The dimensioning of large engineering structural systems (bridges, windmills, offshore structures, ships, airplanes etc.) is always supported on mathematical models or empirical rules that describe the relationship between the exciting forces and the responses everywhere in the structure. Typically, the dimensioning is made on the basis of rules and regulations that have been calibrated to a desired safety level with reference to stochastic modeling of load and resistance variables. The cost of such large engineering structures is high and the impact of a fatal accident of the structure on the owner and society is significant or catastrophic. For that reason it is common to continuously monitor such societal expensive structures and in this way assure that the structure is not overloaded and that it continuously satisfies the safety requirements.

For a modern ship the engine room is equipped with several hundred sensors helping the operator not to overload the engines or other equipment. Almost all components that can be monitored are monitored. The hull, on the other hand, has often only a few sensors that monitor stresses and accelerations at selected locations. These sensors guide the captain during loading and unloading and during voyage to decide on speed reduction or changing course. Similarly, the wings of a windmill will typically be equipped with a small number of strain gauges.

The information from these sensors has in the past only to a limited extent been used in a consistent updating of the actual safety level of the structure and in forecasting the occurrence of future potential adverse events. The use of inspection results on offshore platforms to update the fatigue failure probability is an example of a simple, though effective, use of “monitored” information.

From an engineering point of view it is important that all available significant knowledge about the structure is taken into consideration for the optimal operation of the structure. This knowledge is represented by the mathematical structural model and the stochastic description of the load and response parameters. Being possible by the development of fast information technology, it is attractive currently to use the information about the available stochastic realization of the structure and its load processes as it is embedded in the set of simultaneously monitored data series. This information gathering conveniently takes

the form of consistent Bayesian updating of the prior probability and spectral distributions that were used as the basis for the design of the structure.

It is illustrated herein how the observed signals of the stochastic responses can be used for maximum likelihood estimation of spectral parameters of stationary Gaussian vector processes. Even though the formulated likelihood function in general may turn out to be too complicated to be useful for direct posterior distribution modeling, its curvature properties in the neighborhood of its the maximum point can be calculated numerically. This will allow for approximate statistical uncertainty evaluation by use of the asymptotic large sample Gaussian approximation of the posterior distribution.

The next section describes the superior frame perspective for the use of the described procedure. Section 3 presents the general theory for formulating the likelihood function. The maximum likelihood estimation of parameters in the so-called gamma wave spectrum is illustrated to work well for a simulated realization of the corresponding scalar Gaussian process. In another example a pair of cross-correlated Gaussian processes with the same gamma spectrum is considered. The cross-correlation is given in the form of a specific coherence function family with a single parameter subject to maximum likelihood estimation.

Even though it is required that the considered processes are Gaussian for the likelihood function to be exact, the general property of asymptotic robustness of the estimates with respect to deviations from normality may make the formulated method applicable for slightly non-Gaussian processes. Of course, the method can be applied for all processes that by marginal transformation can be made Gaussian, that is, for the so-called translation processes (Grigoriu 2002).

## 2 STRUCTURAL HEALTH MONITORING

Most structures that are associated with large failure costs are equipped with sensors that register the time variation of strains or accelerations at different locations of the structure. However, these sensors are limited in number and directly from these it is therefore only possible to get information about the local health state of the structure. Therefore, to obtain more complete information of the health state of the structure, it is necessary to make use of some mathematical modeling to describe the relationship between the time varying load and the response. If a full description of the time varying load is available, then it is a reasonably simple task to calibrate the mathematical model such that agreement between the model predictions and observations of the response quantities is obtained. Such a model provides the information for making a detailed health state monitoring of the entire structure.

The actual time varying loading on the structure is rarely monitored in detail. Typically load observations are only obtained as time series of wind speeds or wave elevation measured at a few points. Sometimes the observation may be accompanied by short term estimates of mean wind speed or significant wave height with associated wave period. Although monitoring of the response also is limited to specific locations on the structure it carries information of both the structure properties and the spectrum of its load process.

The generally available procedure for identifying the underlying mathematical model is known as system identification. Basically, system identification is about modeling the relationship between inputs and outputs. The theory for system identification in engineering is to a large extent developed by statisticians and computer scientists within the field of data mining. Within that field the measured signals are typically divided into input $\mathbf{u}(t)$ and output $\mathbf{y}(t)$, see e.g. (Ljung 1999). The relationship in terms of a linear model is

$$\mathbf{y}(t) = \mathbf{F}(\mathbf{p})\mathbf{u}(t) + \mathbf{G}(\mathbf{p})\mathbf{e}(t) \tag{1}$$

where $\mathbf{e}(t)$ is a vector of disturbances and $\mathbf{F}(\mathbf{p})$, $\mathbf{G}(\mathbf{p})$ are given linear operators that depend on a set $\mathbf{p}$ of unknown parameters. System identification is about estimating $\mathbf{p}$. The operators $\mathbf{F}$ and $\mathbf{G}$ can be parameterized to different degrees of detailing that determines the dimension of the parameter vector $\mathbf{p}$. Often different variants of ARMA models are applied (ARMA = Auto Regressive Moving Average). An example in structural engineering is about identifying structural behavior changes caused by fractured steel members (Andersen 1997). The capability of the ARMA-approach for structural system identification is well illustrated in this reference.

The family of stationary ARMA processes of a given order corresponds to a family of mathematical spectra with a number of parameters given by the order of the ARMA process. These parameters are together with the system model parameters $\mathbf{p}$ estimated in some way by comparison with the observed process realizations, be it realizations of the excitation processes and/or the response processes. The maximum likelihood estimation, demonstrated in this paper, delivers a rational algorithm for this estimation, and it is applicable for any family of parametric spectral shapes, of which the ARMA process generated spectra are just one possibility. Moreover it is possible by the Bayesian statistical concept to make use of the stochastic modeling that (in principle) was used for the design of the structure.

The main challenge in making health state monitoring is to use the local and limited information of responses and loading to obtain updated descriptions of load spectra and global structural parameters. By

combining the collected information about load spectra and global parameters with the structural model, updated response spectra are obtained at any location of the structure. This information can be used in a more accurate prediction of local fatigue damage in the structure, for instance. If the probability of fatigue failure at this location is judged to be too high, then the local failure probability can be updated by a local inspection.

By use of Bayesian network graphics (Jensen 2001, Hugin Experts 2002) Fig. 1 illustrates the health state monitoring set up. The white nodes are the mathematical nodes, the light gray (pink) nodes represent monitored data and the darker gray nodes (blue) the uncertain structural parameters and the uncertain load spectral parameters. These are exemplified as the ocean wave spectral parameters $H_s$ (significant wave height), $T_z$ (mean zero level crossing period), and $\xi, \zeta$ (spectral shape parameters). The figure illustrates that monitoring the sea state obviously updates solely the parameters of the load spectra, while monitoring the responses may update both the load spectra and the uncertain structural parameters symbolized by $A$ and $B$. The likelihood principle that can be used for this updating is presented in Section 3.3. Fig. 2 sets the frame for updating a cross-correlation parameter $\alpha$ from monitored simultaneous realizations of two Gaussian processes. The likelihood function for this updating is derived in Section 3.5.

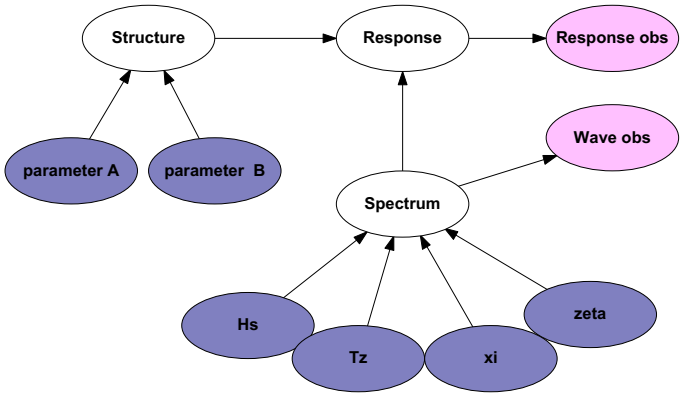


Figure 1. Bayesian network that illustrates how observations may update both the parameters of the load spectra and the global uncertain structural parameters. White nodes are mathematical nodes, the light gray nodes (pink) represent observations and the darker gray nodes (blue) the uncertain structural parameters and uncertain spectral parameters.

Bayesian networks are useful for administration of the operation of complicated technical systems. It gives a clarifying overview and aids the mental process of adding new modeling causal elements to a complicated system. Behind the graphical display of the network is a consistent Bayesian probability calculation program. The Bayesian data updating acts on the probabilistic information stored at the parameter nodes backwards in the system from the data observation nodes through the mathematical system modeling nodes to the parameter nodes. Prediction flows in the opposite direction.

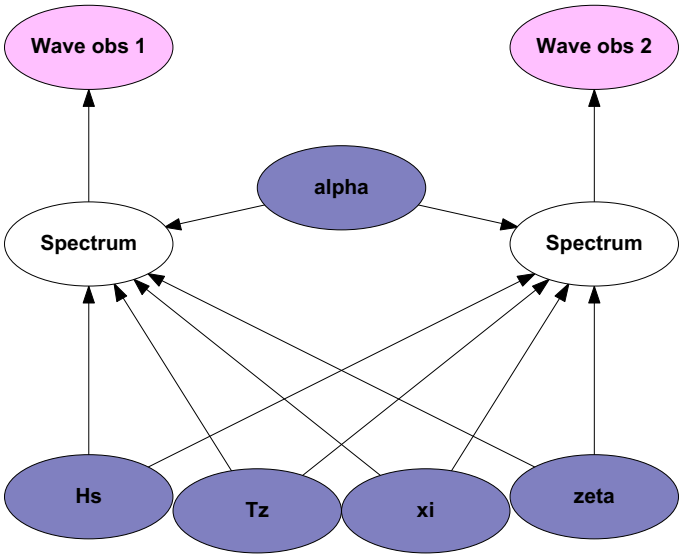


Figure 2. Bayesian network that illustrates how observations may update cross correlated processes.

One purpose of monitoring an operative system is to currently detect if inconsistencies between the stored information and the incoming information show up. It is a sign of inconsistency if the position of the posterior density of the parameters calculated solely from the incoming data and unquestioned prior information deviates too much from the stored density to be characterized as a random deviation. A likelihood ratio test [or within the Bayesian concept: a posterior density ratio test, see e.g. Remark 11.2 in (Ditlevsen & Madsen 2002)] can be applied for this purpose by setting a suitable significance level. If there is no significant deviation from the stored information, the new information can be used to update the stored posterior distribution of the parameters, that is, to decrease the level of the statistical uncertainty.

The numerical search for the maximum likelihood estimates $\hat{\boldsymbol{\mu}}$ (or maximum posterior distribution estimates) of the parameters $\boldsymbol{\mu}$ can be designed to include a calculation of the Hessian $\mathbf{H}$ of the negative logarithm to the likelihood function (or the posterior density) at its maximum. Large sample statistics theory indicates that the likelihood function for a suitable representation of the parameters with sufficient accuracy for practical purposes can be approximated by the normal density $f_{\mathbf{M}}(\boldsymbol{\mu}) \propto \exp[-(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})^{\mathsf{T}}\mathbf{H}(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})/2]$. This simplifies the updating since $\exp[-(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}_1)^{\mathsf{T}}\mathbf{H}_1(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}_1)/2]\exp[-(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}_2)^{\mathsf{T}}\mathbf{H}_2(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}}_2)/2] \propto \exp[-(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})^{\mathsf{T}}(\mathbf{H}_1+\mathbf{H}_2)(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})/2]$ where $\hat{\boldsymbol{\mu}} = (\mathbf{H}_1+\mathbf{H}_2)^{-1}(\mathbf{H}_1\hat{\boldsymbol{\mu}}_1+\mathbf{H}_2\hat{\boldsymbol{\mu}}_2)$ is the vector of updated maximum likelihood estimates.

In the following the paper concentrates on the formulation of the spectral parameter likelihood function with example demonstrations on simulated data.

## 3 MAXIMUM LIKELIHOOD ESTIMATION

### 3.1 Complex representation of Gaussian process

It is general knowledge that a stationary Gaussian random process can be modeled in complex form by

$$Z(t) + i\,\hat{Z}(t) = \int_0^\infty e^{-i\omega t}\,\mathrm{d}W(\omega) \tag{2}$$

where $\mathrm{d}W(\omega)$ is the random complex increment spectral distribution function and $Z(t)$, $\hat{Z}(t)$ are the real and imaginary part, respectively, see e.g. (Cramér & Leadbetter 1967). The random complex increment $\mathrm{d}W(\omega)$ is defined as

$$\mathrm{d}W(\omega) = \mathrm{d}W_\mathrm{r}(\omega) + i\,\mathrm{d}W_\mathrm{i}(\omega) \tag{3}$$

where $\mathrm{d}W_\mathrm{r}(\omega)$ and $\mathrm{d}W_\mathrm{i}(\omega)$ are independent zero mean Gaussian increments along the $\omega$-axis of two independent and identically distributed real Gaussian processes $W_\mathrm{r}(\omega)$ and $W_\mathrm{i}(\omega)$, respectively. The common variance of the increments is

$$E[\,\mathrm{d}W_\mathrm{r}(\omega)^2] = E[\,\mathrm{d}W_\mathrm{i}(\omega)^2] = S(\omega)\,\mathrm{d}\omega \tag{4}$$

and

$$E[\,\mathrm{d}W_\mathrm{r}(\omega_1)\,\mathrm{d}W_\mathrm{r}(\omega_2)] = E[\,\mathrm{d}W_\mathrm{i}(\omega_1)\,\mathrm{d}W_\mathrm{i}(\omega_2)] =$$

$$\delta(\omega_1 - \omega_2)S(\omega_2)\,\mathrm{d}\omega_1\,\mathrm{d}\omega_2 \tag{5}$$

where $\delta(\cdot)$ is Dirac's delta function, and $S(\omega)$ is the spectral density function. Thus for the complex increment $\mathrm{d}W(\omega)$ we have

$$E[\,\mathrm{d}W(\omega_1)\,\mathrm{d}W^*(\omega_2)] =$$

$$E[(\,\mathrm{d}W_\mathrm{r}(\omega_1) + i\,\mathrm{d}W_\mathrm{i}(\omega_1))(\,\mathrm{d}W_\mathrm{r}(\omega_2) - i\,\mathrm{d}W_\mathrm{i}(\omega_2)] =$$

$$2\delta(\omega_1 - \omega_2)S(\omega_2)\,\mathrm{d}\omega_1\,\mathrm{d}\omega_2 \tag{6}$$

$$E[\,\mathrm{d}W(\omega)\,\mathrm{d}W^*(\omega)] = E[|\,\mathrm{d}W(\omega)|^2] = 2S(\omega)\,\mathrm{d}\omega \tag{7}$$

where $*$ means complex conjugate. The real and imaginary parts of (2) are

$$Z(t) = \int_0^\infty [\cos\omega t\,\mathrm{d}W_\mathrm{r}(\omega) + \sin\omega t\,\mathrm{d}W_\mathrm{i}(\omega)]\,\mathrm{d}\omega \tag{8}$$

$$\hat{Z}(t) = \int_0^\infty [-\sin\omega t\,\mathrm{d}W_\mathrm{r}(\omega) + \cos\omega t\,\mathrm{d}W_\mathrm{i}(\omega)]\,\mathrm{d}\omega \tag{9}$$

from which it is seen that $\hat{Z}(t)$ is the Hilbert transform of $Z(t)$. The common covariance function of $Z(t)$ and $\hat{Z}(t)$ is

$$\mathrm{Cov}[Z(0), Z(t)] = \mathrm{Cov}[\hat{Z}(0), \hat{Z}(t)] =$$

$$\int_0^\infty S(\omega)\cos\omega t\,\mathrm{d}\omega \tag{10}$$

To be able to operate numerically with the Gaussian random process $Z(t)$ it is replaced by the trigonometric polynomial

$$Q(t) = \sum_{k=1}^{N} [\Delta W_\mathrm{r}(\omega_k)\,\cos\omega_k t + \Delta W_\mathrm{i}(\omega_k)\,\sin\omega_k t] \tag{11}$$

where $\omega_k = k\Delta\omega$, $\Delta W_\mathrm{r}(\omega_k)$ and $\Delta W_\mathrm{i}(\omega_k)$ for all $k = 1, \ldots, N$ are mutually independent Gaussian random variables of zero mean and standard deviation $\sqrt{S(\omega_k)\Delta\omega}$. This corresponds to the discretization $0 < \Delta\omega < 2\Delta\omega < \ldots < N\Delta\omega$. The discretization is reasonably chosen such that numerical integration of the spectrum $S(\omega)$ over the interval $[0, N\Delta\omega]$ by the trapezoidal formula, say, and the chosen discretization deviates less than some chosen small value from the exact variance $\int_0^\infty S(\omega)\,\mathrm{d}\omega$.

Since the functions $\cos\omega_k t$, $\sin\omega_k t$, $k = 1, \ldots, N$, define an orthogonal function system over the interval $[0, T]$, where $T = 2\pi/\Delta\omega$ is the period of the trigonometric polynomial (11), it follows directly that

$$\Delta W_\mathrm{r}(\omega_k) = \frac{2}{T}\int_0^T Q(\tau)\cos\omega_k\,\tau\,\mathrm{d}\tau$$

$$\Delta W_\mathrm{i}(\omega_k) = \frac{2}{T}\int_0^T Q(\tau)\sin\omega_k\,\tau\,\mathrm{d}\tau \tag{12}$$

### 3.2 Nonparametric maximum likelihood estimation of spectrum

Assume now that a sample $q(t)$ of $Q(t)$ has been observed over the time interval $[0, T]$. By substituting $q(t)$ for $Q(t)$ in (12) the corresponding realizations $\Delta w_\mathrm{r}(\omega_k)$ and $\Delta w_\mathrm{i}(\omega_k)$ of the $2N$ random variables $\Delta W_\mathrm{r}(\omega_k)$ and $\Delta W_\mathrm{i}(\omega_k)$ are obtained. Estimates of the $N$ variances $S(\omega_k)\Delta\omega$ are obtained by maximizing the likelihood function

$$\prod_{k=1}^{N}\frac{1}{S(\omega_k)\Delta\omega}\varphi\left(\frac{\Delta w_\mathrm{r}(\omega_k)}{\sqrt{S(\omega_k)\Delta\omega}}\right)\varphi\left(\frac{\Delta w_\mathrm{i}(\omega_k)}{\sqrt{S(\omega_k)\Delta\omega}}\right) \propto$$

$$\prod_{k=1}^{N}\frac{1}{S(\omega_k)\Delta\omega}\exp\left(-\frac{\Delta w_\mathrm{r}(\omega_k)^2 + \Delta w_\mathrm{i}(\omega_k)^2}{2S(\omega_k)\Delta\omega}\right) \tag{13}$$

which without restrictions on $S(\omega)$ takes its maximum for

$$S(\omega_k) = \frac{\Delta w_\mathrm{r}(\omega_k)^2 + \Delta w_\mathrm{i}(\omega_k)^2}{2\Delta\omega} =$$

$$\frac{T}{4\pi}\left[\left(\frac{2}{T}\int_0^T q(\tau)\cos\omega_k\tau\,\mathrm{d}\tau\right)^2 +\right.$$

$$\left.\left(\frac{2}{T}\int_0^T q(\tau)\sin\omega_k\tau\,\mathrm{d}\tau\right)^2\right] = \frac{1}{\pi T}\left|\int_0^T q(\tau)e^{i\omega_k\tau}\,\mathrm{d}\tau\right|^2 \tag{14}$$

for $k = 1, \ldots, N$. The estimate fluctuates to a certain degree along the $\omega$-axis, and the more so the shorter is the time $T$. Usually some smoothness restrictions are imposed to reflect that $S(\omega)$ has analytical properties such as differentiability to some order. Herein we will assume that a spectral family $S(\omega; \boldsymbol{\mu})$ is given a priori, where $\boldsymbol{\mu}$ is a set of parameters to be estimated from the sample $q(t)$, possibly in combination with prior knowledge about $\boldsymbol{\mu}$.

### 3.3 Maximum likelihood estimation of spectral parameters

For the given spectral family $S(\omega; \boldsymbol{\mu})$ the likelihood function $L(\boldsymbol{\mu}|q)$ is obtained from (13) and (14) as

$$L(\boldsymbol{\mu}|q) \propto \tag{15}$$

$$\prod_{k=1}^{N} \frac{1}{S(\omega_k; \boldsymbol{\mu})} \exp\left( \frac{-1}{\pi T S(\omega_k; \boldsymbol{\mu})} \left| \int_0^T q(\tau) e^{i\omega_k \tau} \, \mathrm{d}\tau \right|^2 \right)$$

The log-likelihood function is then

$$-\log L(\boldsymbol{\mu}|q) + \text{constant} = \sum_{k=1}^{N} \log S(\omega_k; \boldsymbol{\mu}) +$$

$$\frac{1}{\pi T} \sum_{k=1}^{N} \frac{1}{S(\omega_k; \boldsymbol{\mu})} \left| \int_0^T q(\tau) e^{i\omega_k \tau} \, \mathrm{d}\tau \right|^2 \approx$$

$$\frac{1}{\Delta\omega} \Bigg( \int_0^{N\Delta\omega} \log S(\omega; \boldsymbol{\mu}) \, \mathrm{d}\omega +$$

$$\frac{1}{\pi T} \int_0^{N\Delta\omega} \frac{1}{S(\omega; \boldsymbol{\mu})} \left| \int_0^T q(\tau) e^{i\omega\tau} \, \mathrm{d}\tau \right|^2 \mathrm{d}\omega \Bigg) \tag{16}$$

for a sufficiently small increment $\Delta\omega = 2\pi/T$. Thus we have approximately

$$L(\boldsymbol{\mu}|q) \propto \exp\Bigg( -\frac{T}{2\pi} \int_0^{\omega_{\max}} \log S(\omega; \boldsymbol{\mu}) \, \mathrm{d}\omega -$$

$$\frac{1}{2\pi^2} \int_0^{\omega_{\max}} \frac{1}{S(\omega; \boldsymbol{\mu})} \left| \int_0^T q(\tau) e^{i\omega\tau} \, \mathrm{d}\tau \right|^2 \mathrm{d}\omega \Bigg) \tag{17}$$

where $\omega_{\max} = N\Delta\omega$.

The example illustrations below are made for the so-called gamma spectrum

$$S(\omega \,|\, \xi, \zeta, H_s, T_z) = A\omega^{-\xi} e^{-B\omega^{-\zeta}}, \ \omega > 0 \tag{18}$$

where

$$A = a(H_s/4)^2 (2\pi/T_z)^{\xi-1}, \ B = b(2\pi/T_z)^{\zeta} \tag{19}$$

$$a = \zeta \Gamma[(\xi-1)/\zeta]^{\frac{\xi-3}{2}} \Gamma[(\xi-3)/\zeta]^{-\frac{\xi-1}{2}} \tag{20}$$

$$b = \Gamma[(\xi-1)/\zeta]^{\frac{\zeta}{2}} \Gamma[(\xi-3)/\zeta]^{-\frac{\zeta}{2}} \tag{21}$$

where $\Gamma(\cdot)$ is the gamma function. Thus we have the four parameters $H_s > 0$, $T_z > 0$, $\xi > 3$, $\zeta > 0$ in the parameter vector $\boldsymbol{\mu}$. For the particular parameter values $\xi = 5$ and $\zeta = 4$ this spectrum is also known as the Pierson-Moskowitz spectrum applied for random ocean waves with the interpretation of $H_s$ and $T_z$ as the significant wave height and the mean zero crossing period, respectively, of a stationary sea state.

### 3.4 Estimation example 1

It is seen from the likelihood function expression (17) that there will be numerical problems when trying to estimate a spectrum that comes close to zero within the integration interval $[0, \omega_{\max}]$. This is the case for the gamma spectrum. However, the problem is easily circumvented by adding a small amount of simulated white noise of given spectral intensity $\epsilon$ to the available realization of the random process. This adds the positive constant $\epsilon$ to the spectrum that therefore still depends solely on the unknown parameters of the original spectrum.Thus no information is lost by the addition of the noise, and the numerical problems are avoided.

Fig. 3 top shows a white noise modified gamma spectrum corresponding to the parameters $\xi = 5$ and $\zeta = 4$ (the Pierson-Moskowitz spectrum), $\omega_{\max} = 4$ $\mathrm{s}^{-1}$, and with white noise intensity $\epsilon = 0.05\ \mathrm{m}^2\mathrm{s}$. The sea state is given by the parameter values $(H_s, T_z) = (4\,\mathrm{m}, 7\,\mathrm{s})$. The middle diagram shows a period of a realization of a Gaussian trigonometric polynomial defined by discretization of the spectrum by a frequency step of $\Delta\omega = 0.01\ \mathrm{s}^{-1}$. The period then becomes $T = 2\pi/\Delta\omega = 200\pi$. The bottom diagram shows the corresponding non-parametric maximum likelihood estimate (14) of the spectrum. Fig. 4 shows a set of contour curves for the likelihood function (17) corresponding to the simulated realization in Fig. 3. It is seen that the parameter values used in the simulations are reasonably well estimated by the maximum points of the restricted likelihood functions given $(H_s, T_z)$ (top four diagrams) or given $(\xi, \zeta)$ (bottom four diagrams).

The unrestricted maximum likelihood estimates are calculated for the first two simulated realizations to $(\xi, \zeta, H_s, T_z)$ = (5.18, 4.04, 3.71 m, 6.94 s), and $(\xi, \zeta, H_s, T_z)$ = (5.14, 4.25, 3.95 m, 6.28 s). These estimates are quite well reproducing the values $(\xi, \zeta, H_s, T_z)$ = (5, 4, 4 m, 7 s) used for the simulations.

Interpreting the likelihood function as a posterior density, it is seen that $(\xi - 3, \zeta)$ tend to be situated on a hyperbolic curve, that is, the product $(\xi - 3)\zeta$ is seen to be much less uncertain than each of the two parameters are. This invites to consider the pair of logarithms $[\log(\xi - 3), \log\zeta]$ as parameters in stead of $(\xi - 3, \zeta)$, also noting that the last parameters must both be positive. Then the pair $[\log(\xi - 3), \log\zeta]$

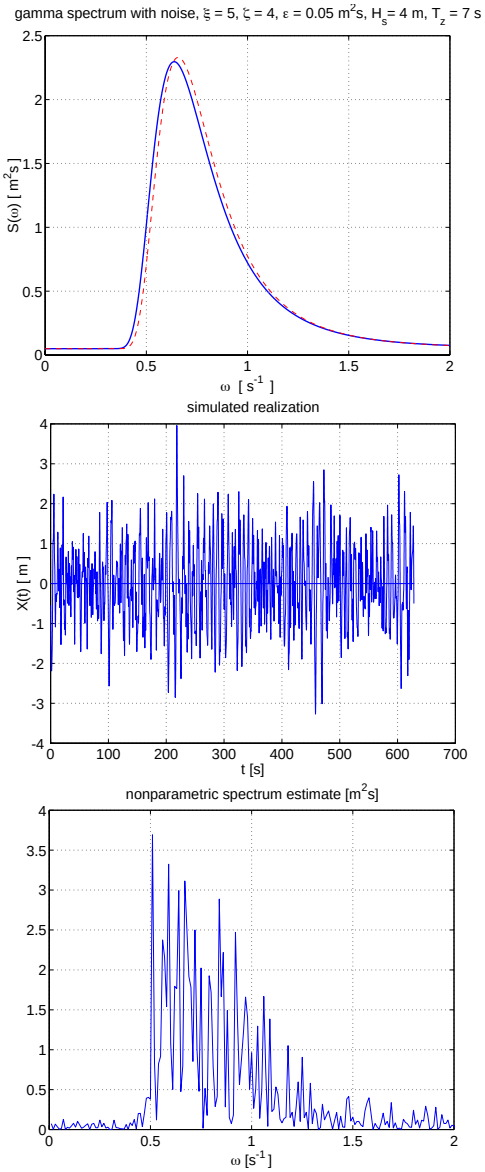


Figure 3. Top: Gamma spectrum with added white noise spectral of intensity $\epsilon = 0.05$ m$^2$s plotted for $\omega \in [0, \omega_{\max}/2]$ (full curve; the dashed curve is defined in the text). Middle: Realization of Gaussian process according to this spectrum. Bottom: Nonparametric maximum likelihood estimates of the spectral values according to (14).

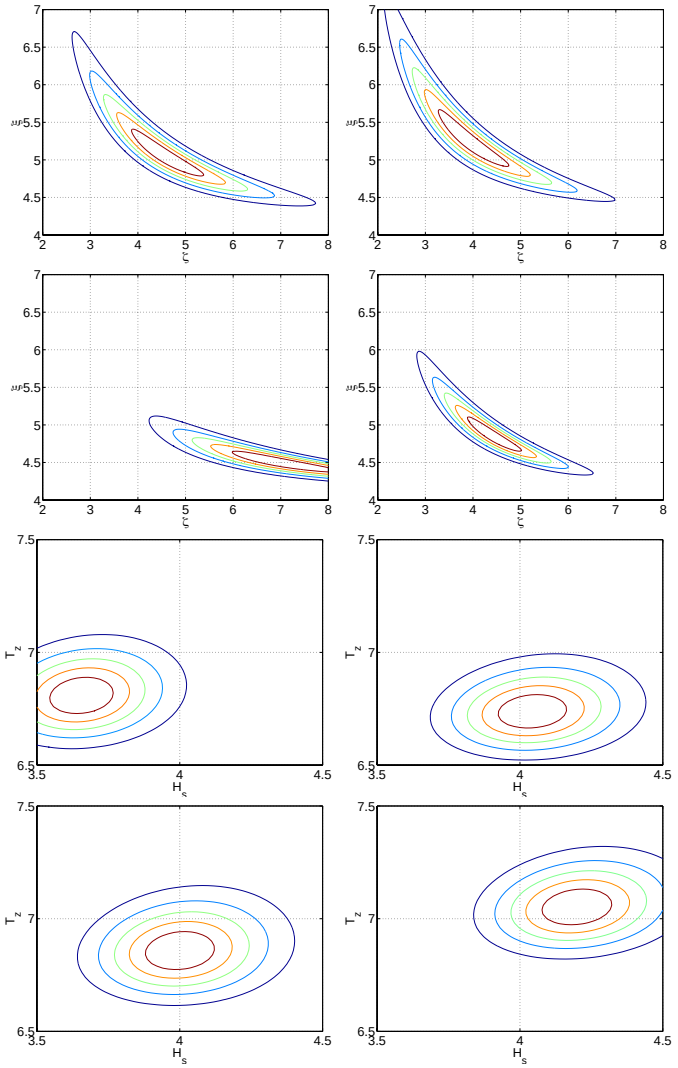


Figure 4. Likelihood function contour plots for $(\zeta, \xi)$ given $(H_s, T_z) = (4\text{m}, 7\text{s})$ (first four plots) and for $(H_s, T_z)$ given $(\zeta, \xi) = (4, 5)$ (last four plots). The first plot of each kind corresponds to the simulated realization in Fig 3, while the other plots correspond other independent realizations.

tends to be closely situated along a straight line of negative slope. A normal distribution approximation will obviously be better after this logarithmic transformation than before. The normal distribution is defined with mean at the maximum likelihood estimate point for $[\log(\xi - 3), \log\zeta]$ and with inverse covariance matrix equal to the Hessian of the negative log-likelihood function at the same point. This makes updating operational following the rule described at the end of Section 2. If the logarithmic transformation is not made, the updating of the parameter estimates must, in principle, be through multiplication of the likelihood functions followed by search for the maximum of the product. Contrary to what is the case for normal distribution factors, the maximum point of the product is not a convex linear combination of the two maximum points of the factors. This is caused by the hyperbolic bend of the two factors.

An example with simulation of 1400 sample realizations gave the average point $(5.16, 4.20)$ of the 1400 maximum likelihood estimates of $(\xi, \zeta)$ and empirical standard deviations $(0.83, 1.25)/\sqrt{1400} = (0.02, 0.03)$ (and correlation coefficient $-0.72$) of the averages. These numbers clearly indicate that the averages are biased estimates of the parameters known to be $(\xi, \zeta) = (5, 4)$. The bias is to the positive side as it should be expected from the bend of the likelihood functions. In Fig. 3 (top) the dotted curve shows the spectrum corresponding to the average estimates $(\xi, \zeta, H_s, T_z) = (5.16, 4.20, 4.00\text{ m}, 6.95\text{ s})$ from the 1400 realizations. It is seen that the effect of the bias is modest. However, even though the bias is small, it may be worth to reduce it by the logarithmic transformation. After logarithmic transformation the averages of the 1400 values of $\log(\xi - 3)$ and $\log\zeta$ are 0.703 and 1.393, respectively, giving the estimates $(\xi, \zeta) \approx (5.02, 4.03)$. Since $H_s$ and $T_z$ cannot be negative either, it may be worth also to replace these by

their logarithms. In fact, the 1400 size sample indicates slight bias to the negative side of the average of the maximum likelihood estimates of $T_z$.

### 3.5 Estimation from correlated records

Let $\mathbf{Q}(t)$ be a vector of random trigonometric polynomials $Q_1(t), \ldots, Q_n(t)$ as (11). For each of these the coefficients are given by (12). The covariance matrix of the $n$ coefficients of $\cos\omega_k t$ or to $\sin\omega_k t$ is

$$\mathbf{S}(\omega_k)\Delta\omega = \{S_{rs}(\omega_k)\}_{nn}\Delta\omega \tag{22}$$

where $S_{rs}(\omega)$ is the cross-spectrum (or auto-spectrum if $r = s$) between $Q_r(t)$ and $Q_s(t)$. The likelihood function (15) becomes replaced by

$$L(\boldsymbol{\mu}|\mathbf{q}) \propto \prod_{k=1}^{N} \frac{1}{\det[\mathbf{S}(\omega_k;\boldsymbol{\mu})]} \exp\Big[-\frac{2}{T^2\Delta\omega}$$

$$\Big(\int_0^T \mathbf{q}(\tau)^{\mathsf{T}} \cos\omega_k\tau \,\mathrm{d}\tau\; \mathbf{S}^{-1}(\omega_k;\boldsymbol{\mu}) \int_0^T \mathbf{q}(\tau)\cos\omega_k\tau\,\mathrm{d}\tau +$$

$$\int_0^T \mathbf{q}(\tau)^{\mathsf{T}} \sin\omega_k\tau \,\mathrm{d}\tau\; \mathbf{S}^{-1}(\omega_k;\boldsymbol{\mu}) \int_0^T \mathbf{q}(\tau)\sin\omega_k\tau\,\mathrm{d}\tau\Big)\Big] \tag{23}$$

which can be approximated by

$$L(\boldsymbol{\mu}|\mathbf{q}) \propto \tag{24}$$

$$\exp\Big[-\frac{T}{2\pi}\int_0^{\omega_{\max}} \log[\det\mathbf{S}(\omega;\boldsymbol{\mu})]\,\mathrm{d}\omega - \frac{1}{2\pi^2}\int_0^{\omega_{\max}}$$

$$\Big(\int_0^T \mathbf{q}(\tau)^{\mathsf{T}} \cos\omega\tau \,\mathrm{d}\tau\; \mathbf{S}^{-1}(\omega;\boldsymbol{\mu}) \int_0^T \mathbf{q}(\tau)\cos\omega\tau\,\mathrm{d}\tau +$$

$$\int_0^T \mathbf{q}(\tau)^{\mathsf{T}} \sin\omega\tau \,\mathrm{d}\tau\; \mathbf{S}^{-1}(\omega;\boldsymbol{\mu}) \int_0^T \mathbf{q}(\tau)\sin\omega\tau\,\mathrm{d}\tau\Big)\,\mathrm{d}\omega\Big]$$

for a sufficiently small increment $\Delta\omega = 2\pi/T$.

### 3.6 Estimation example 2

To test the efficacy of parameter estimation based on the likelihood function (24), realizations of a pair of cross-correlated processes $[X_1(t), X_2(t)]$ are generated. The two processes are identically distributed and defined as the trigonometric polynomial considered above generated for the gamma spectrum with the parameter values $(\xi, \zeta, H_s, T_z) = (5, 4, 4\,\mathrm{m}, 7\,\mathrm{s})$. To have cross-correlation between the two component processes a coherence function of the form $\exp(-\alpha\omega)$ is chosen, and the aim is to demonstrate the estimation of the coherence parameter $\alpha$. As before white noise of arbitrary small intensity $\epsilon = 0.05$ is added to the

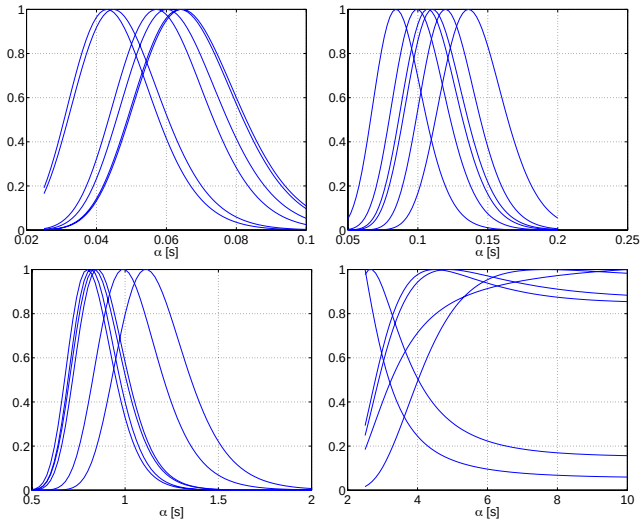


Figure 5. Normalized likelihood functions of the coherence parameter $\alpha$ for given $(\zeta, \xi) = (4, 5)$ and $(H_s, T_z) = (4\,\mathrm{m}, 7\,\mathrm{s})$ defined by (24) and obtained from the simulated realizations of a pair of processes $[X_1(t), X_2(t)]$ with coherence function $\exp(-\alpha\omega)$. The four examples are $\alpha = 0.05, 0.1, 1, 5$ (top left to bottom right) with six simulated realizations in each example.

processes. This means that the two Gaussian coefficients in each of the two trigonometric polynomials corresponding to the frequency $\omega_k$ get the variances $[S(\omega_k) + \epsilon]\Delta\omega$ and that the coefficients to $\cos\omega_k\tau$ or to $\sin\omega_k\tau$ in the two processes are correlated with the covariance $S(\omega_k)\exp(-\alpha\omega_k)\Delta\omega$. The simulation is for each pair of coefficients made by simulating realizations of three independent standard Gaussian variables $Z, Z_1, Z_2$ and setting

$$X_i = \sqrt{\frac{S(\omega_k)e^{-\alpha\omega_k}}{S(\omega_k)+\epsilon}}\, Z + \sqrt{1 - \frac{S(\omega_k)e^{-\alpha\omega_k}}{S(\omega_k)+\epsilon}}\, Z_i \tag{25}$$

for $i = 1, 2$. Thereafter $X_1$ and $X_2$ are multiplied by the standard deviation $\sqrt{[S(\omega_k) + \epsilon]\Delta\omega}$.

The spectral matrix for substitution into the likelihood function (24) thus is

$$\mathbf{S}(\omega;\boldsymbol{\mu},\alpha) = S(\omega;\boldsymbol{\mu})\begin{bmatrix} 1 & e^{-\alpha\omega t} \\ e^{-\alpha\omega t} & 1 \end{bmatrix} + \epsilon\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{26}$$

where $\boldsymbol{\mu} = (\xi, \zeta, H_s, T_z)$ are the parameters of the gamma spectrum $S(\omega;\boldsymbol{\mu})$.

Fig. 5 shows the results of maximum likelihood estimation of the coherence parameter $\alpha$ from 6 independent simulations in each of the cases $\alpha = 0.05, 0.10, 1.0, 5.0$. For making the comparison easier the likelihood functions are normalized by division with the maximal value. It is seen that the estimation is good for the smaller values of $\alpha$, but becomes more diffuse if $\alpha$ is as large as 5. This is to be expected when considering that the coherence decreases increasingly faster to zero with $\omega$ for increasing $\alpha$. It is well known that it is difficult to estimate correlation coefficients close to zero with great accuracy.

## 4 LIKELIHOOD RATIO TEST

The knowledge of the likelihood function makes it possible to set up a test procedure for detecting changes of the incoming monitored data as compared to the stored information. Let $L_1(\boldsymbol{\mu}_1)$ be the likelihood function corresponding to the stored information and let $L_2(\boldsymbol{\mu}_2)$ be the likelihood function corresponding to the new data. Then the ratio

$$R = \frac{\max_{\boldsymbol{\mu}}\{L_1(\boldsymbol{\mu})L_2(\boldsymbol{\mu})\}}{\max_{\boldsymbol{\mu}_1,\boldsymbol{\mu}_2}\{L_1(\boldsymbol{\mu}_1)L_2(\boldsymbol{\mu}_2)\}} \tag{27}$$

is a well-known indicator for changes [Likelihood-ratio test (Neyman & Pearson 1928)]. In the statistics of identically distributed independent observations, the random variable $-2\log R$ is asymptotically, as the sample size grows beyond limits, distributed as the $\chi^2$-distribution with $m-k$ degrees of freedom. The numbers $m$ and $k$ are the dimensions of the total unrestricted parameter space and of the parameter hypothesis subspace, respectively. In the present case $m$ is twice $k$ where $k$ is the number of parameters in $\boldsymbol{\mu}$. Thus the test of the zero hypothesis of no change is to check whether the realization of $-2\log R$ is smaller than some chosen upper tail fractile value in the $\chi^2$-distribution with $m-k$ degrees of freedom. For the spectral parameter likelihood functions considered here it turns out that the asymptotic $\chi^2$-distribution is an excellent approximation. Fig. 6 shows a comparison between the theoretical $\chi^2$-distribution of 4 degrees of freedom and the empirical distribution function obtained by simulating 700 independent pairs of realizations corresponding to the spectrum in Fig. 3.

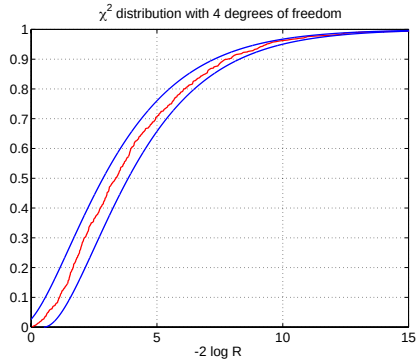


Figure 6. The empirical distribution of $-2\log R$ corresponding to a simulated sample of 700 pairs of independent process realizations compared to the theoretical asymptotic $\chi^2$-distribution with 4 degrees of freedom shifted $\pm 0.5$ along the abscissa axis.

## 5 CONCLUSIONS

Spectral parameter estimation based on the maximum likelihood principle is demonstrated to work well. The likelihood function is formulated in explicit form on the basis of the process of independent Gaussian increments that defines the given stationary Gaussian process. The likelihood function is only amenable to numerical calculations to find its maximum point. Singularities in the likelihood function may disturb the numeric analysis, but the problem is easily removed by adding a simulated white noise of small intensity to the observed process realization.

The knowledge of the likelihood function opens the possibility of making rational Bayesian updating during active monitoring of a structure subjected to process loads. Possibly the well known robustness of estimation based on the normal distribution with respect to deviations from normality makes the procedure applicable in practical information monitoring also for non-Gaussian processes that do not deviate too radically from Gaussian processes.

It is envisaged that significant changes of the parameters for new monitored data can be detected by a likelihood ratio test and thus be taken as warnings about damage in progress or other malfunctioning of the monitored structure. If the test is passed, the new data can be utilized for updating of the stored likelihood function information about the structure and its loads. This updating can with sufficient accuracy be made by normal distribution approximation to the likelihood function for suitably defined parameters. The normal distribution is defined by its covariance matrix which is the inverse of the Hessian matrix of the negative logarithm to the likelihood function calculated at the point of maximum likelihood, and it is centered at the maximum likelihood estimates.

## REFERENCES


Andersen, P. (1997). *Identification of Civil Engineering Structures using Vector ARMA Models*. Phd thesis, Department of Building Technology and Structural Engineering, Aalborg University, Denmark, Sohnholmgaardsvej 57, DK-9000 Aalborg, Denmark. www.civil.auc.dk/ i6pa/thesis.htm.

Cramér, H. & M. R. Leadbetter (1967). *Stationary and Related Stochastic Processes*. New York: Wiley.

Ditlevsen, O. & H. O. Madsen (1996, 2002). *Structural Reliability Methods*, Edition 1. Wiley-Interscience-Europe, Chichester. Edition 2.1 (internet edition, 2002): www.mek.dtu.dk/staff/od/books.htm.

Grigoriu, M. (2002). *Stochastic Calculus*. Birkhäuser, Boston, Basel, Berlin.

Hugin Experts (2002). *Hugin Professional, v.6.1*. Hugin Experts A/S, Niels Jernes Vej 10, DK 9220 Aalborg, Denmark. www.hugin.com.

Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*. Springer Verlag, New York.

Ljung, L. (1999). *System Identification – Theory for the User* (2nd ed.). PTR Prentice Hall, Saddle River, N.J.

Neyman, J. & E. S. Pearson (1928). On the use and interpretation of certain criteria for the purposes of statistical inference. *Biometrika* 20A:175 and 263.