Working Papers in Statistics
No 2016:7

Department of Statistics
School of Economics and Management
Lund University

Distributions of spatial wave size for random fields

KRZYSZTOF PODGÓRSKI, LUND UNIVERSITY

IGOR RYCHLIK, CHALMERS UNIVERSITY OF TECHNOLOGY

Distributions of spatial wave size for random fields

Krzysztof Podgórski
Department of Statistics
Lund University
Email: Krzysztof.Podgorski@stat.lu.se

Igor Rychlik
Mathematical Sciences
Chalmers University of Technology
Email: rychlik@chalmers.se

***Abstract*—A method of measuring three-dimensional spatial wave size is proposed and statistical distributions of the size characteristics are derived in explicit integral forms for Gaussian sea surfaces. New definitions of wave characteristics such as the crest-height, the length, the size and the wave front location are provided in fully dimensional context. The joint statistical distributions of these wave characteristics are derived using the Rices formulas for expected numbers of local maximum and distance from a local maximum to a level crossing countour. Review of the Rice's method to study crossing distributions will be given.**

I. Introduction

In oceanography and marine technology the sea elevation seen as a three dimensional moving surface is perceived as propagating trains of apparent waves. Determination of statistical properties of three dimensional wave sizes is of great importance in ocean-engineering applications such as a design of mechanical structures that have dimensions of the same order of magnitude as the wave, e.g., oil platforms, wave energy harvesters, or even floating airports, bridges, large ships and piers. The importance of developing methodology to address spatial properties of waves lies in the fact that the impact of a wave on a structure depends on the wave shape in space. In this paper we present an extended approach that captures better three-dimensional geometry of spatial waves. This approach is based on a new definition of wave sizes and on the generalized Rice formula that facilitates derivation of their statistical distributions.

The distributions describe variability of sea waves properties for a stationary Gaussian sea defined by the directional power spectrum associate with the instantaneous wave climate at the location. Then the long-term distributions of some wave characteristics from those simulations are obtained by averaging the distributions over various spectra related to the wave climate variability. Simulation approach is possible, however the computation cost of such a procedure would be prohibitive, especially if one wants good accuracy for high - and thus rare - crests. In this paper we demonstrate a method of direct estimation of the stationary sea distributions using the probabilistic model of W for a given directional spectrum. Through this one can avoid simulations of large fields that are required to statistically analyze rare extreme events.

The paper is organized as follows. Section II contains introductory material: notation, definitions of wave characteristics in one dimensional records, and a short review of distributions of wave characteristics in Gaussian seas. The new findings, definitions of wave characteristics of waves in sea surface, are given in Section III. A short discussion of methods that are used for evaluation of statistical distribution of the proposed waves characteristics is also given in this section. Validation of the proposed methodology and some examples are presented in Section IV. Conclusions and an extensive list of references follows.

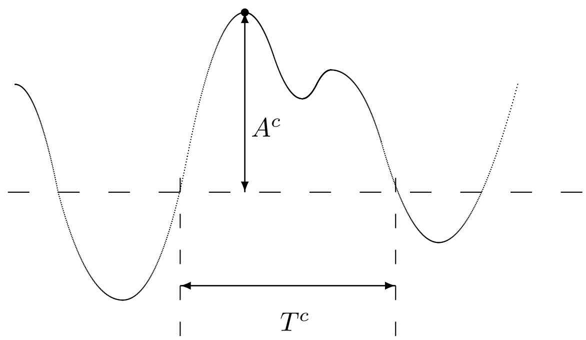
II. Distributions of waves characteristics along one dimension

The sea surface elevation is often measured at a fixed location and saved in the form of a time series. Here, we use time context although the measurements can be also made linearly in the space in which the case some obvious modification of terminology is needed. The so called zero-crossing wave is used to define apparent waves. In this approach wave is a part of a record between two consecutive upcrossings of the mean sea level. For simplicity it is assumed that the latter, also called the still water level, is set to zero. Important wave characteristics are crest height A^c and crest duration (half period of a wave) T^c, see Figure 1*(Top)*. The crest A^c of a wave is the maximum value between a zero-upcrossing and the following zero-downcrossing. Similarly, the crest duration T^c is the time distance between a zero-upcrossing and the following downcrossing.

Since spatial records are seldom available the spatial wave characteristics relays on mathematical modeling of sea surface dynamics. A simple model, that is used here, assumes that each sinus wave component travel independently at its own celerity. It can be shown that it results in a probabilistic assumption that states that $W(x, y, t)$ is a Gaussian field, see [1] for details. For a long-crested sea for which $W(x, y, t) \approx W(x, t)$ one often assumes that sinus waves travels with negative velocity and hence a zero upcrossings is chosen to mark the front of an apparent wave.

Space wave characteristics are defined in spatial signals $W(x) = W(x, 0, 0)$ and $W(y) = W(0, y, 0)$. In the signals, for example $W(x)$, one can identify the spatial apparent waves which are characterized by its crest A_x^c and crest length L_x^c which is the distance between upcrossing and the following downcrossings of the zero level. The characteristics A_y^c, L_y^c are defined in a similar way for waves in the spatial signal $W(y)$. In the following for simplicity of notation we will write A, L_x and L_y for crest height and crest lengths in x, y, directions, respectively.

Distributions of wave characteristics describe variability of the observed waves during stationary sea conditions. These are understood as limits of histograms, empirical cumulative distributions etc of wave characteristics, as observation time/region grow without bounds. The limiting distributions can depend on



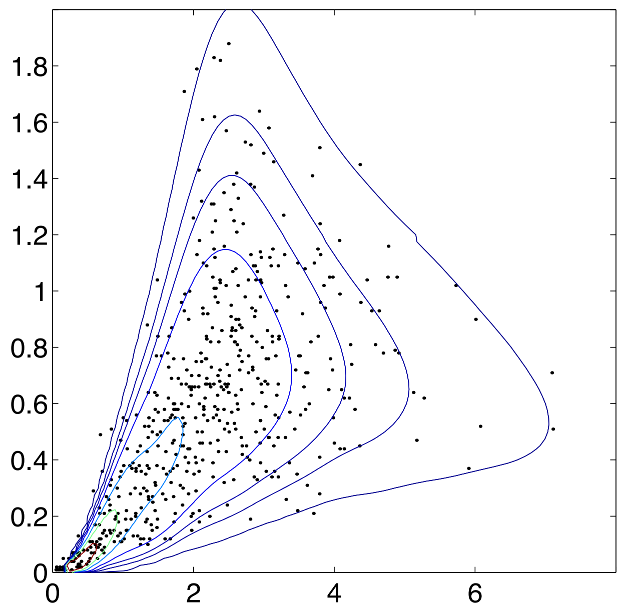


Fig. 1. *Top:* The definition of wave wave crest A^c (crest amplitude) and half period (crest length) T^c. *Bottomt:* The distribution of crest amplitude and length computed based on the Rice formula and the actual values observed in records of 534 temporal waves at shallow water of the coast of Africa.

the considered populations of waves, i.e. the way waves are collected (counted) and on particular definition of an apparent wave characteristics.

In Figure 1*(Bottom)*, we have illustrated the approach by showing the isolines of the distribution for the crest amplitude and length computed using a popular class of spectra in ocean engineering the JONSWAP spectrum (JOint North Sea WAve Project) with the parameters fit to the empirical records. This distribution is compared with the empirical distribution of the amplitudes and lengths observed in the data (dots). One can notice that among 534 waves used in this example only very few are observed in the area of the extreme crest amplitudes and lengths. In contrast, the distribution based on Rice formula provides quite accurate density of the distribution also in regions with very few observations. This data set and the numerical routines for obtaining the figure can be found conveniently in the WAFO toolbox, [2].

Finding a joint probability density for crest height A^c and length T^c is difficult since it requires derivation of three dimensional density of local maximum and locations of the nearest zeros of the sea, see for [3] for details. Finding distributions $P(T^c > t)$, and $P(A^c \leq h | T^c = t)$ are much easier tasks as can be seen in (6) and (7).

In this paper we are focusing on properties of sea waves characteristics. However the derived results can have applications in other sciences. For example computations of $P(A^c > h)$ is also important in studying risk for failures due to fatigue process in metals, see Remark 1

The concept of zero crossing waves, presented in Figure 1*(Top)*, is difficult to generalize to the spatial waves (or spatio-temporal) case. An alternative definition of waves in space and time has been proposed in [4]. In that work the so called "acting wave crest" was introduced. Here we shall slightly modify this approach to cover waves in time, space or time and space.

Definition 1: For a time point t let $\mathcal{D}$ be the largest interval centered at t such that $W(s) > 0$ for all s lying in the interior of the interval $\mathcal{D}$. Point t is a center of a wave if $W(t) = \max_{s\in\mathcal{D}} W(s)$. The crest height of the wave is defined as $A = W(t)$ while the length of the interval $\mathcal{D}$, T, say, is the length of the crest.

The definition applied to the signal in Figure 1*(Top)* will give crest amplitude $A = A^c$ however $T < T^c$. In general applying Definition 1 to a measured signal $W(t)$ may result in larger number of waves.

A. *Distribution of wave characteristics - Rice's method*

For completeness we present expressions for the distributions of some wave characteristics obtained by means of various generalizations of Rice formula. The classical form of the latter yields an explicit form for the intensity of zeros in Gaussian process $W(t)$ having power spectral density $S(\omega)$. If we restrict only to the upcrossings of the zero level, the average intensity (the average number of the uppcrossings per the time unit) is given by

$$\mu = \frac{1}{2\pi}\sqrt{\lambda_2/\lambda_0}. \tag{1}$$

Here λ_i's are spectral moments of $S(\omega)$. The formula can be reinterpreted as the intensity of waves defined by the zero upcrossing method. Despite that (1) is well known, let us mention a slightly more general version that extends its validity beyond the Gaussian case, i.e. one considers a zero mean stationary process $W(t)$ for which the expected number of times $W(t) = u$ in a region Λ, $N_\Lambda(u)$, say, is given by

$$E[N_\Lambda(u)] = \|\Lambda\| \int |z| f_{\dot{W}(0),W(0)}(z,u)\,dz.$$

Clearly, for the Gaussian case and due to independence between $\dot{W}(0)$ and $W(0)$, it turns to the classical Rice's formula originally presented in [5], [6] and [7]:

$$E[N_\Lambda(u)] = \frac{\|\Lambda\|}{\pi}\sqrt{\frac{\lambda_2}{\lambda_0}}e^{-u^2/2\lambda_0}.$$

Hence $\frac{1}{\pi}\sqrt{\frac{\lambda_2}{\lambda_0}}$ is the intensity of zeros, solutions of $W(t) = 0$. Since there are two zeros per wave yielding the formula for the intensity of waves (1). In what follows we present integral formulas used to derive the distributions of wave characteristics. We comment on the principles used to derive the formulas but we avoid mathematical technicalities and refer to other work for details.

B. Counting marked waves along lines

As mentioned above a “global” property that each second zero of $W(t)$ is an zero-upcrossing which marks front of a wave combined with Rice's formula gave the intensity of wave in (1). Finding distributions of wave characteristics is a harder problem since it requires accounting for some suitable local properties of $W(t)$ around a time t such that $W(t) = 0$. We will illustrate this “local” approach by providing a simple direct argument for the intensity of waves μ given (1).

Following the definition of waves presented in Figure 1, any zero upcrossing is a front of a wave. Obviously, if $W(t) = 0$ and

$$B = \text{“derivative } \dot{W}(t) \text{ is positive”} \tag{2}$$

is true, then at location t is a front of a wave. Another (more complex) example of local properties of $W(t)$ is

$$B_r = \text{“derivative } \dot{W}(t) \text{ is positive and for all } s \in [0, r] \text{ and } W(t+s) \geq 0\text{”.} \tag{3}$$

Again, if $W(t) = 0$ and B_r is true, then location t is a front of a wave having half-period longer than r. The expected values of zeros of W such that a condition ”B” (or ”B_r”) holds can be computed using the following version of Rice's formula

Generalized Rice Formula: *If for a stationary Gaussian process W its marked zeros, i.e. times $W(t) = 0$ such that condition B holds, are homogeneously spread in time, then the expected number of marked zeros in Λ, $N_\Lambda(B)$ is given by*

$$E[N_\Lambda(B)] = \|\Lambda\| \int P(B|\dot{W}(0) = z, W(0) = 0) \cdot |z| f_{\dot{W}(0),W(0)}(z, 0)\, dz, \tag{4}$$

see Lemma 7.5.2 in [8] for more details.

Combining (4) with (2-3) one obtains formulas for intensity of waves μ and intensity of waves with half-period exceeding r, $\mu(T^c \geq r)$, respectively, viz.

$$\mu = \int_0^{+\infty} z f_{\dot{W}(0),W(0)}(z, 0)\, dz = \frac{1}{2\pi}\sqrt{\frac{\lambda_2}{\lambda_0}} \tag{5}$$

$$\mu(T^c \geq r) = \int_0^{+\infty} P(B_r|\dot{W}(0) = z, W(0) = 0)) \cdot |z| f_{\dot{W}(0),W(0)}(z, 0)\, dz. \tag{6}$$

Note that derivation of (5) requires some simple calculus and employing property that $W(0)$ and $\dot{W}(0)$ are independent. (One also uses assumption that sea has zero mean.) Intensity $\mu(T^c \geq r)$ in (6) has to be evaluated using numerical integration. A function RIND from toolbox WAFO could be used for this purpose.

The method can be used to compute fairly complex events. For example, in [3], the joint density of the distance from the crossing to the crest S^c and the half-amplitude A^c conditionally on the value of the half-period $T^c = t$, was obtained

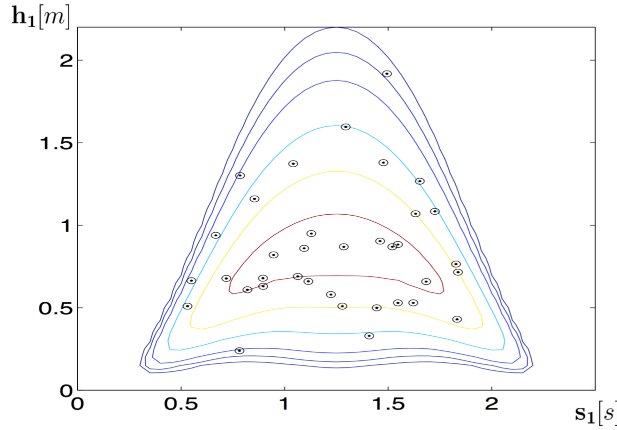


Fig. 2. The conditional probability density function of the crest height A^c and its distance S^c from the level crossing given $T^c = 2.5s$ computed from the spectrum estimated from real data vs. the actual records.

$$f_{S^c,A^c|T^c=t}(s, h) = \frac{\int_{\mathbb{R}^3} |uvz| P(t, h, u, v, z) f(0, 0, h, 0, u, v, z)\; du\; dv\; dz}{\mu \cdot f_{T^c}(t)}, \tag{7}$$

where $P(t, h, u, v, z)$ is the conditional probability of $B_{th} = \{0 < W(u) \leq h,$ for all $u \in (0, t)\}$ given that seven dimensional vector

$$(W(0), W(t), W(s), \dot{W}(s), \dot{W}(0), \dot{W}(t), \ddot{W}(s))$$

takes value $(0, 0, h, 0, u, v, z)$ while $f(\mathbf{x})$ is the density of this vector.

Remark 1: In this paper we are presenting means to derive probability distributions of sea waves characteristics. However the derived expressions can also be applied to problems in other sciences. Here we present an important application taken from mechanics.

For example, let $W(t)$ be a zero mean stationary stress. For a signal $u + W(t)$, which now has mean u, evaluate the intensity $\mu(A^c > h)$ and denote it by $\mu^{osc}(u, h)$, the so called oscillation intensity or the intensity of upcrossings of interval $[u, u+h]$, see [9]. Knowing the distribution of local maximums and the oscillation intensity one can derive the rainflow cycles distribution, see [10] for definition of rainflow cycles. This distribution is then used to predict the fatigue life prediction of metallic components.

III. Geometry of waves in space

In this section a method of counting waves is introduced and then used to define waves characteristics for the sea surface $W(x, y)$ that is located in three dimensions.

A stationary Gaussian sea is fully characterized by a directional spectrum $S(\omega, \alpha)$ which describe energy of waves moving along a line having angle α with x axis. A zero mean Gaussian sea having directional spectrum $S(\omega, \alpha)$ can be approximated (with arbitrary accuracy) by a sum of independent cosine waves with Rayleigh distributed amplitudes and

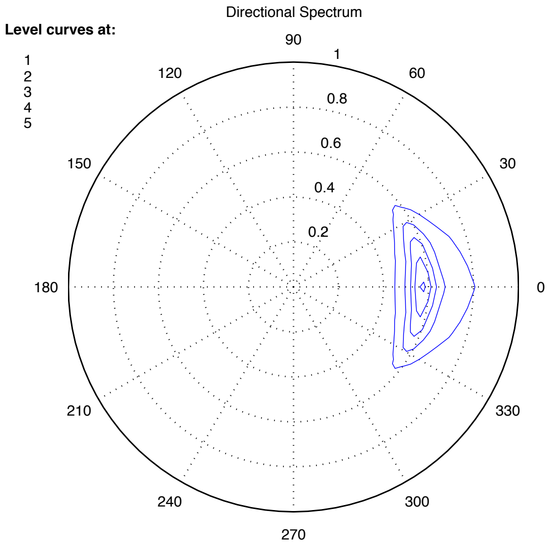


Fig. 3. The directional spectrum used in the examples.

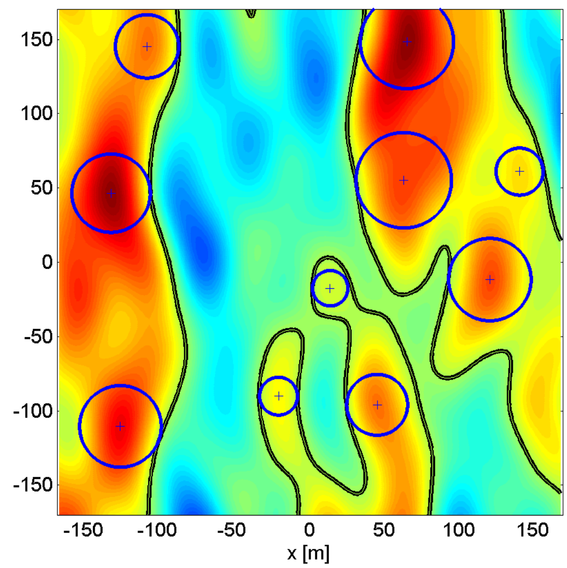


Fig. 4. Illustration of wave definition given in Definition 2. The crosses mark the local maxima.

uniformly distributed phases. For the directional spectrum let $S_{jk} = S(\omega_k, \alpha_j)$ and for $\alpha_j \in [0, 2\pi], \omega_k > 0$ define the field

$$\begin{aligned} W(x,y,t) \quad &\approx \quad \sum \sqrt{S_{jk}} R_{jk} \sqrt{d\omega_k d\alpha_j} \qquad (8) \\ &\cdot \cos(t\,\omega_k - \frac{\omega_k^2}{g}(x\cos\alpha_j + y\sin\alpha_j) + \phi_{jk}), \end{aligned}$$

where R_{jk} are independent standard Rayleigh variables, ϕ_{jk} are independent uniformly distributed over $[0, 2\pi)$ phases that are also independent of R_{jk}, and $(d\omega_k, d\alpha_j)$ are infinitesimally small increments of the rectangular grid over arguments of the spectrum (the accuracy of approximation depends on how fine is the grid).

The methodology will by illustrated for waves defined in Gaussian sea $W(x, y)$ having directional power spectrum shown in Figure 3. The spectrum has following parameters; significant wave height $h_s = 7$ [m], average crest length in x and y directions $E[L_x], E[L_y]$, defined in (10), equal to 21π [m] and 58π [m], respectively. In fact, the spectrum is a directional spread of a member from the JONSWAP parametric family of sea spectra with the half-time period equal to $E[T^c] = 4.7$[s].

A. *Defining waves in space*

The following definition of wave is a generalization of Definition 1.

Definition 2: For a point $\mathbf{p}$ let $\mathcal{D}$ be the largest disk centered at $\mathbf{p}$ such that $W(\mathbf{q}) > 0$ for all $\mathbf{q}$ lying in the interior of the disk. The point $\mathbf{p}$ is a center of a wave if $W(\mathbf{p}) = \max_{\mathbf{q}\in\mathcal{D}} W(\mathbf{q})$, see Figure 4 for illustration.

The crest height of the wave is defined as $A = W(\mathbf{p})$ while the radius R of $\mathcal{D}$ defines the half-length of the crest. The point $\mathbf{q}$ where the disk $\mathcal{D}$ is tangential to the zero level is a front of the wave. Further the angle Θ such that $\mathbf{q} = \mathbf{p} + R(\cos\Theta, \sin\Theta)$ will be called the direction of the front of the wave (with respect to the location of the crest).

Note that the Definition 2 applied to a stationary sea, would locate waves homogeneously on (x, y)-plane. This is convenient when statistical properties of waves observed over a large area are of interest.

Remark 2: Many images consist of two or more phases, where a phase is a collection of homogeneous zones. For example, the phases may represent the presence of different sulphides in an ore sample. Frequently, these phases exhibit very little structure, though all connected components of a given phase may be similar in some sense. As a consequence, random set models are commonly used to model such images The wave characteristics introduced in Definition 2 could be applied to analyze such images as is sketched below.

Again consider a surface $u + W(x, y)$ and denote by A_u, R_u, Θ_u, The marginal distributions of the characteristics depend on the threshold u and are two dimensional functions describing variability of the surface. Actually in [11], [12], this kind of variability measures have been used for segmentation of images containing three types of ore. The distributions of gray scales for the ore were quit similar making segmentation difficult. However the ore had different degree of hardness making variability of gray scales different. The variability were described using the oscillation intensity similar to $\mu(A_u > h)$.

B. *Size of a wave*

The sizes of disks $\mathcal{D}$, given in Definition 2, could serve as a proxy of the spacial spread of an apparent wave. However real waves are seldom isotropic, see Figure 4, and hence disk maybe inadequately describe the spread of a wave. Instead, we propose to use appropriate ellipses instead of circles and to introduce them we need to define the normalized sea surface.

1) The normalized sea surface: Consider a stationary sea with directional spectrum $S(\omega, \alpha)$. The normalized sea surface $\widetilde{W}$ is obtained from W by scaling it in both arguments and in value so that $\widetilde{W}$ has variance and variances of its partial

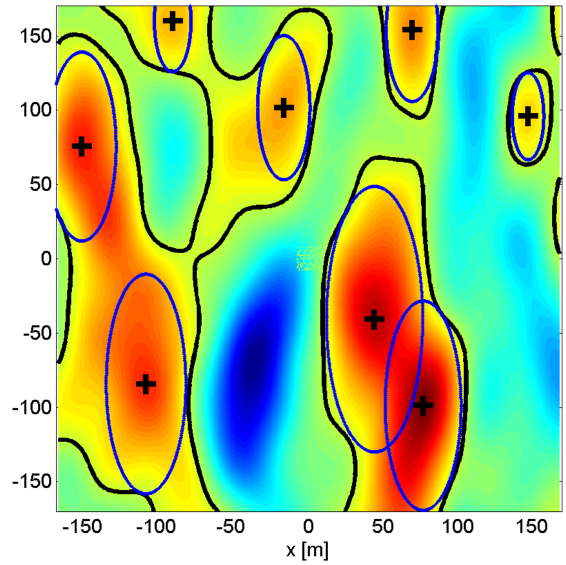


Fig. 5. Size of waves in $W(x, y)$ by applying Definition 2 to the normalized field $\widetilde{W}(x, y)$.

derivatives equal to one. This scaling has natural interpretation in terms of the normalized directional wave lengths as discussed next. It is defined using spectral moments λ_{ij} that are given by

$$\lambda_{ij} = \int_0^{2\pi} \int_0^{+\infty} S(\omega, \alpha) \left(\frac{\omega^2}{g}\right)^{i+j} (\cos\alpha)^i (\sin\alpha)^j \, d\omega \, d\alpha. \tag{9}$$

The significant wave height is then $h_s = 4\sqrt{\lambda_{00}}$ while

$$E[L_x] = \pi\sqrt{\lambda_{00}/\lambda_{20}} \qquad E[L_y] = \pi\sqrt{\lambda_{00}/\lambda_{02}} \tag{10}$$

are the average half wavelengths in x, y directions, respectively. With this notation and terminology the normalized field can be written explicitely

$$\widetilde{W}(x, y) = W\big(xE[L_x]/\pi, yE[L_y]/\pi\big)/\sqrt{\lambda_{00}}. \tag{11}$$

Thus the normalized sea has the half-lengths of the zero crossings waves both in the direction x and the direction y equal to π.

2) Size of a wave - area enclosed by an ellipse: Suppose that, for a sea state, the directional spectrum is known. In such a case for a given sea surface $W(x, y)$ one can rescale it to obtain the normalized field $\widetilde{W}(x, y)$, defined in (11). Now using Definition 2 one can identify waves in $\widetilde{W}$. The crest position and half-length of the ith wave is denoted by $\widetilde{p}_i$, $\widetilde{R}_i$, respectively. The circles with centers at $\widetilde{p}_i$ and radius $\widetilde{R}_i$ become ellipses in the original (x, y) coordinates, see Figure 5 for illustration. The size of ith wave can be gauged by the area enclosed by the ith ellipse, viz.

$$S_i = \pi\, L_x\, L_y \widetilde{R}_i^{\,2}. \tag{12}$$

C. Intensity of waves

In evaluation of safety of maritime operation, “dangerous” (potentially harmful) events are often defined using wave characteristics, e.g crest height, period (length), wave steepness etc. Estimates of risks involves computations of frequency of the so identified “dangerous” waves. In what follows we define precisely what is meant by frequencies of the waves. We start with some necessary notation.

Let consider stationary (homogeneous) sea which will be observed in a region Λ, say. In the region one will count waves. Denote by N_Λ number of waves found in Λ and let the number of waves for which an event B happens be denoted by $N_\Lambda(B)$. Waves and “dangerous” waves (as described by an occurrence of B for such a wave) are most often spread homogeneously over the surface. Consequently there is a constant $\mu(B)$, called intensity, such that the expected number of “dangerous” waves in Λ is equal the intensity of waves times size of Λ, viz.

$$E[N_\Lambda(B)] = \mu(B) \cdot \|\Lambda\|,$$

where $\|\Lambda\|$ is the size of Λ (length, area, volume, etc.). Obviously the intensity of waves μ, say, is equal to $\mu(B)$ for an event B which is always true. Finally, the probability of B is defined by

$$P(B) = \mu(B)/\mu. \tag{13}$$

The probability can be formally interpreted, for the ergodic seas, as a limit of an empirical frequency $N_\Lambda(B)/N_\Lambda$ as the size $\|\Lambda\|$ increases without bound.

For a given sea state intensity of waves μ is easy to estimate and is often included as one of sea state parameters. Hence, by (13), quantities $\mu(B)$ and $P(B)$ could be equivalently used. However if estimates of μ and $P(B)$ are taken from different sources one has to be cautious that the same definitions of waves were used when estimating the parameters.

In many applications intensities $\mu(B)$ of interest are very small comparing to the duration of a sea state. Consequently the intensities have to be extrapolated from the available data. Statistics and probability theory are a natural frameworks for such extrapolations. Another approach is to assume a stochastic model for sea surface variability and then compute the intensities using suitable probabilistic tools. One is often using a Gaussian model for the sea. In such a case the power spectral density (psd) of a sea state defines the model and one can (at least in principle) evaluate distributions of various wave characteristics and use those to estimate $\mu(B)$. In particular Rice's formula [5], [6] and [7] and its generalizations, see e.g. in [13], [14] and references therein, are very useful tools to evaluate $\mu(B)$.

Dynamics of waves will not be considered here and we refer to [15] and [16] where the dynamics of sea wave was statistically described, e.g. a probability distribution of velocities of wave crests in $W(x, y, t)$ were presented.

D. Counting waves at a sea surface

In this section will consider similar problems as in Section II-B although this time a sea surface $W(x, y)$ will be considered instead of a record $W(t)$ in time or along a line. The height of local maximum will be denoted by M. First formulas for the intensity of local maximums M higher than h, i.e. $\mu(M > h)$, will be given. The problems were first studied by Longuet-Higgins in [15].

For a local maximum located at $\mathbf{p} = (x, y)$ exceeding level h if for the sea surface gradient

$$(W_x(\mathbf{p}), W_y(\mathbf{p})) = (0, 0),$$

and the statement

$$B = \text{"matrix } \ddot{W}(\mathbf{p}) \text{ is negative definite and } W(\mathbf{p}) > h\text{"}$$

is true. The intensity $\mu(M > h)$ can be evaluated using the following version of Rice's formula.

Multivalued Rice Formula: *Consider a multivalued random functions $\mathbf{X}(\mathbf{p}) : R^n \to R^n$ such that at the point $\mathbf{p}$ a statement B about $\mathbf{X}$, its derivatives or any other random process is true. If zeros satisfying condition B are spread homogeneously in space then, under some conditions see e.g. [13], [17], [14], the following equation holds*

$$\mu(B) = \int P(B|\dot{\mathbf{X}}(\mathbf{0}) = \dot{\mathbf{x}}, \mathbf{X}(\mathbf{0}) = \mathbf{0}) \, |\det \dot{\mathbf{x}}| f_{\dot{\mathbf{X}}(\mathbf{0}),\mathbf{X}(\mathbf{0})}(\dot{\mathbf{x}}, \mathbf{0}) \, d\dot{\mathbf{x}}. \quad (14)$$

In order to utilize (14) for derivation the intensity of waves with crest above h, let us define $\mathbf{X}(\mathbf{p}) = (W_x(\mathbf{p}), W_y(\mathbf{p}))$ then

$$\mu(M > h) = \int_C P(W(\mathbf{0}) > h|\dot{\mathbf{X}}(\mathbf{0}) = \dot{\mathbf{x}}, \mathbf{X}(\mathbf{0}) = \mathbf{0}) \, |\det \dot{\mathbf{x}}| f_{\dot{\mathbf{X}}(\mathbf{0}),\mathbf{X}(\mathbf{0})}(\dot{\mathbf{x}}, \mathbf{0}) \, \mathbf{d}\dot{\mathbf{x}}, \quad (15)$$

where $C = \{\dot{\mathbf{x}} : \dot{\mathbf{x}} \text{ is negative definite matrix}\}$. The so expressed intensity $\mu(M > h)$ has to be computed numerically and no explicit algebraic formula has been found yet. Even a formula for intensity of local maximums in $W(x, y)$, derived first in [15], is expressed using the Legendre elliptic integrals of the first and second kinds.

By (13) the probability that positive local maximum observed at the sea exceeds h is given by

$$P(M > h) = \mu(M > h)/\mu(M > 0).$$

Differentiating (15) on h gives the pdf of M viz.

$$f_M(h) = c \int_C |\det \dot{\mathbf{x}}| f_{\dot{\mathbf{X}}(\mathbf{0}),\mathbf{X}(\mathbf{0}),W(\mathbf{0}))}(\dot{\mathbf{x}}, \mathbf{0}, \mathbf{h}) \, \mathbf{d}\dot{\mathbf{x}}, \quad (16)$$

where $C = \{\dot{\mathbf{x}} : \dot{\mathbf{x}} \text{ is negative definite matrix}\}$ while $c = 1/\mu(M > 0)$ is the normalization constant. The integral in (16) has to be computed numerically however for high values of h, which is often of the main interest, asymptotic formulas for $f_A(h)$ can be given, see [18], [14] and [19].

1) Joint distribution of crest height and length A, R.: In this section we give formula for intensity of waves having crest height exceeding h and crest half-length longer than r, i.e. $\mu(A > h, R > r)$. Since for stationary sea surface $W(x, y) = W(\mathbf{p})$ waves having crest and length exceeding some fixed thresholds are homogeneously spread in space formula one can use (14) to evaluate $\mu(A > h, R > r)$, viz.

$$\mu(A > h, R > r) = \int_C P(B|\dot{\mathbf{X}}(\mathbf{0}) = \dot{\mathbf{x}}, \mathbf{X}(\mathbf{0}) = \mathbf{0}) \, |\det \dot{\mathbf{x}}| f_{\dot{\mathbf{X}}(\mathbf{0}),\mathbf{X}(\mathbf{0})}(\dot{\mathbf{x}}, \mathbf{0}) \, \mathbf{d}\dot{\mathbf{x}}, (17)$$

where $C = \{\dot{\mathbf{x}} : \dot{\mathbf{x}} \text{ is negative definite matrix}\}$ while

$$B = \text{"}W(\mathbf{0}) > h \text{ and for all } \mathbf{p} \in \mathcal{D}_r,\ W(\mathbf{0}) \geq W(\mathbf{p}) > 0\text{"}.$$

Here $\mathcal{D}_r$ is a disk centered at $\mathbf{0}$ having radius r. The integral in (17) has to be computed numerically. This is not an easy task since B is a function of infinitely many variables $W(\mathbf{p})$, $\mathbf{p} \in \mathcal{D}_r$ such that $W(\mathbf{p}) > 0$, and thus it has to be approximated by suitable discretization. For example, since a sea surface is often observed on a grid $\mathbf{p}_i$, say, then in practice the following is a natural approximation

$$B \approx \text{"}W(\mathbf{0}) > h,\ W(\mathbf{0}) \geq W(\mathbf{p}_i) > 0 \text{ for all } \mathbf{p}_i \in \mathcal{D}_r\text{"}.$$

Using the approximative definition of B, the integral in (17) becomes finite dimensional. Methods to evaluate numerically (14) have been discussed in several papers, see e.g. [20], [21]. The program RIND in WAFO [2] can be used to evaluate (15). Alternatively, one can use MC simulations (for example, Fourier snapshots methods are presented in [4]) to approximate the conditional probabilities $P(B|\cdot)$ and then numerically integrate the integrals in (14).

Derivation of the formula for joint pdf of A, R is quite complex. It involves generalizations of Rice's formula for non-homogeneously spread zeros of multivalued function $\mathbf{X}(\mathbf{p})$. In fact, it helps to introduce an additional variable Θ and derive a formula for the joint density of A, R, Θ, where Θ is a direction from the crest to the nearest point on the zero level contour. We will not give formulas here.

IV. WAVE SIZES DISTRIBUTIONS FOR A LONG-CRESTED SPECTRUM

In this section we illustrate the proposed characteristics of waves in space by presenting the probability density functions (marginal and joint) for; crest height A, crest half-length R, direction of wave front Θ and the wave size S, for waves in a Gaussian sea having directional power spectrum shown in Figure 3. The spectrum is defined by the following parameters; significant wave height $h_s = 7$ [m], average crest length in x and y directions $E[L_x], E[L_y]$, defined in (10), equal to 21π [m] and 58π [m], respectively.

The presentation of the results is organized as follows. Iin Section IV-A, the pdf of (R, Θ), defined according to Definition 2, is computed. Note that the point $(R\cos\Theta, R\sin\Theta)$ is the location of a crest front when the orign of the coordinate system is placed at the location of a wave crest. In other words R is the distance to the closest wave front from a wave crest. Further in Section IV-A computation of the probability that wave crest exceeds a fixed threshold is discussed. Then the joint pdf of (A, R, Θ) is presented in Section IV-B. Finally in Section IV-C we show the pdf of size S of a wave.

A. Joint probability density of R, Θ

The wave characteristics (R, Θ) have been called the half-length of the crest and the direction of the wave front, respectively. The probability distribution of (R, Θ) has the following interpretation. Let place position of a randomly chosen wave crest at origin $(0, 0)$ and find the position $\mathbf{q}$ of the front of the wave. Then $\mathbf{q} = R(\cos\Theta, \sin\Theta)$.

The joint pdf of R, Θ is presented in Figure 6 *(Top)* in the polar coordinate system. The marginal pdfs of R, Θ are

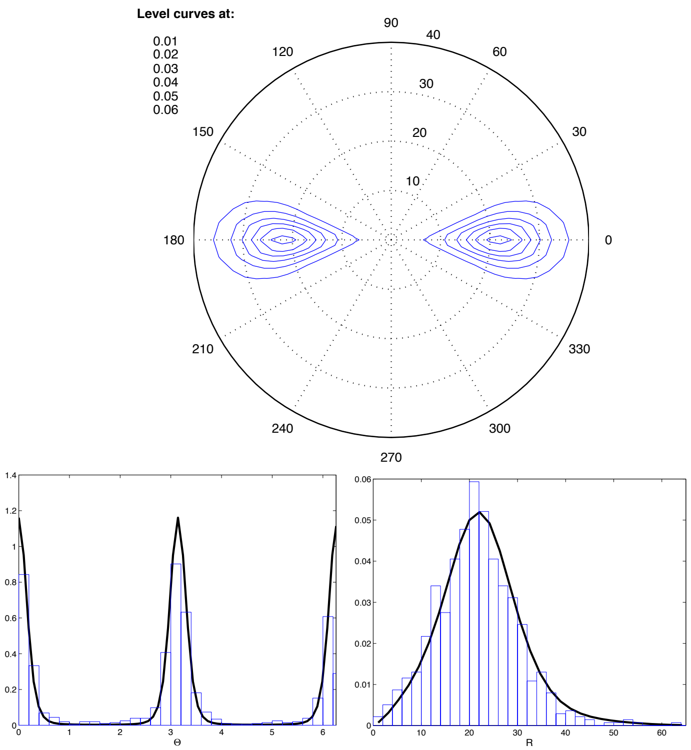


Fig. 6. *Top:* The joint pdf of R, Θ. *Left-Bottom:* the pdf of Θ. *Right-Bottom:* the pdf of R. The marginal distribution are compared with the normalized histogram obtained from simulations.

derived by means of numerical integration of the joint pdf and presented in Figure 6 *(Left-Bottom)*– the azimuth Θ and in *(Right-Bottom)* – the distance R. Both are compared for the accuracy check with the normalized histograms of R and Θ, respectively, found in simulated surfaces $W(x, y)$. One can see that histograms and pdfs $f_R(r)$, $f_\Theta(\theta)$ agree very well.

It agrees with intuition in particularly for nearly long-crested seas.

ACKNOWLEDGMENT

The first author was supported from Riksbankens Jubileumsfond Grant Dnr: P13-1024 and Swedish Research Council Grant Dnr: 2013-5180, while the second was partial support by Swedish Research Council Grant 340-2012-6004 and by Knut and Alice Wallenberg stiftelse.

B. The joint pdf of A, R and Θ.

We turn now to the joint density of crest height A, the distance to zero level contour R, and the angle Θ. We have chosen the region of high waves such that the crest hight A exceeds $0.5\,h_s = 3.5$ meter. The level $0.5h_s$ is chosen to be moderately high in order to allow comparisons with values of A, R, Θ obtained from simulated records. One can roughly estimate that $P(A > 3.5) = 0.16/0.94$, which means that 17% of waves have crest exceeding $0.5\,h_s = 3.5$[m], which is sufficiently high to be see frequently in the records. Consequently, one concludes that in a long-crested sea, the fraction of wave crests above the threshold $0.5h_s$ is less than 13.5%.

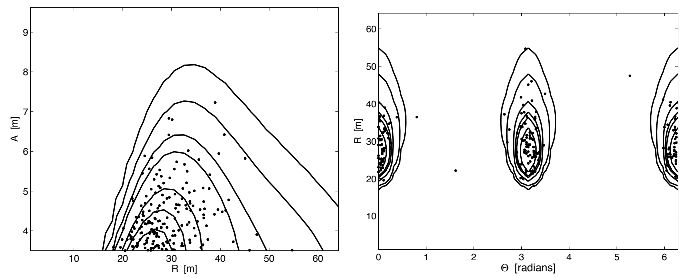


Fig. 7. Conditional distributions given that crest is higher than $0.5\,h_s$ ($h_s = 7$ [m]). *Top:* The conditional joint pdf of the crest height A and the crest half-length R. *Bottom:* The conditional pdf of R and Θ. In the left plot contour lines contain 30%. 50%, 70% , 90%, 95%, 99%, 99.9% of the total probability while the contour lines in the right plot have the corresponding percentage set to 10%, 30%. 50%, 70% , 90%, 95%.

In Figure 7, the marginal densities of R, A (left plot) and Θ, R (right plot) given that $A > 0.5\,h_s$, are presented. The densities are integrals of the joint pdf of A, R, Θ normalized by the probability $P(A > 0.5\,h_s)$. The levels of plotted contour lines are relative so that regions enclosed by the contour lines contain specified fractions of the total probability mass. Such a choice of level lines helps in visual evaluations of the accuracy of estimated joint pdf. More precisely, it enables comparisons with observed values of A, R, Θ shown as dots in the plots, since one can count fractions of dots included in the contours and compare those with the contour level defining fractions.

In order to check the accuracy of the numerical integrations involved in the computations of the densities shown in Figure 7, one thousand of waves were extracted (at random) from simulated surfaces $W(x, y)$. Since the probability that a crest is higher than $0.5\,h_s$ is 0.17 one expects to have 170 waves satisfying the condition $A > 0.5\,h_s$. Actually, there were 183 such waves with crest above the threshold. This exceeds the expected value 170 by about one standard deviation thus the difference is not significant.

For the extracted waves one has evaluated values of h, r and θ. The pairs (r, θ) were plotted as dots in the left plot and similarly pairs (θ, r) are shown as dots in the right plot. The isolines of the conditional pdfs are selected so that in the average there should be $180 \cdot x\%$ dots included in the contour. Fractions $x\%$ are specified in the figure. It is easier to count the point outside a contour line. This yields $0.001 \cdot 170 = 0.17, 1.7, 8.5, 17, 51, 85, \ldots$ expected dots outside the respective contour lines. In the left plot of Figure 7, one finds $0, 1, 9, 21, 61, 91, \ldots$ dots. Similarly in the right plot one expects to have $1.7, 8.5, 17, \ldots$ points outside the isolines while in the right plot $4, 8, 22, \ldots$ are counted. This confirms that the accuracy of the estimated densities is quite good.

Finally, we note that as the height of crest A is getting higher, the evaluation of the joint pdf of A, R and Θ by means of generalized Rice's formula is getting easier (and faster) since for Gaussian fields the field around high local maximum can be approximated by a function with some random coefficients, see [17]. This is in contrast to using the simulation approach to approximate the pdfs involving A by counting frequencies – very high crests occur very seldom.

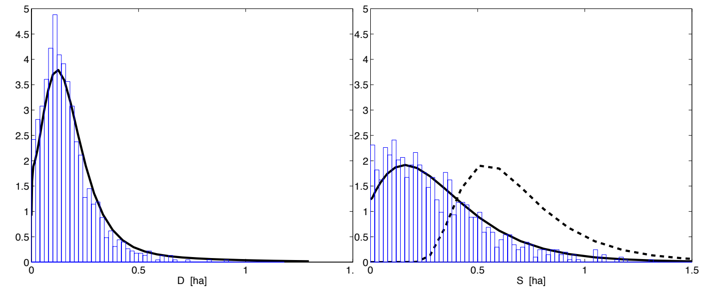


Fig. 8. *Left:* Probability density of D [ha] (18) compared with the normalized histogram based on simulations. *Right:* Comparison of probability densities of S [ha], see (19), with the normalized histogram. The dashed line represents the conditional pdf of S given that the crest height exceeds $0.5\, h_s$.

C. Size of waves S

In Figure 6 *(Bottom)*, the pdf of the crest half-length R is shown. Variable R, defined in Definition 2, is the radius of the disk centered at a wave crest and underneath of the sea surface, see Figure 4. Thus, the distribution of the disk's area D could be used to describe the spatial size of waves. The density of $D = \pi\, R^2$ is given by

$$f_D(s) = \frac{1}{2\sqrt{\pi\, s}} f_R\left(\sqrt{s/\pi}\right). \tag{18}$$

For the considered symmetric directional spectrum, the pdf of D (18) is shown in Figure 8 *(Left)*. The pdf is compared with the normalized histogram of D values extracted from simulations of field W. The plot demonstrates again very good accuracy of computed R-pdf.

It was mentioned in Section III-B that the area enclosed by ellipse S, defined in (12), could be a more adequate proxy describing waves geometry rather than the area of the disk D. In order to find the distribution of S, one can first to evaluate the pdf of crest lengths $\widetilde{R}$ evaluated in the standardized sea surface $\widetilde{W}(x, y)$ defined in (11). The density of S can be then computed by rescaling the disks by the mean half-wave lengths in the directions of the x and y axes via

$$f_S(s) = \frac{1}{2\sqrt{\pi L_x L_y\, s}} f_{\widetilde{R}}\left(\sqrt{s/\pi L_x L_y}\right). \tag{19}$$

The probability density of S is shown in Figure 8 *(Right)*. Comparison of the pdf with the normalized histogram of S confirms good accuracy of evaluated pdf of $\widetilde{R}$. From the plots given in Figure 8 one can conclude that the areas for ellipse S are statistically larger than the areas of disc D. Finally in Figure 8 *(Right)* the dashed line is the conditional pdf of S given that the crest height A exceeds $0.5\, h_s$. Not surprisingly one can observe that high waves also have large sizes.

CONCLUSIONS

A new method to measure size of waves in spatial sea surface records is presented. It is shown that the generalized Rice's formula can be used to evaluate the probability density functions of wave sizes in a Gaussian sea including the joint probability density of the crest height and the length of waves. The pdfs are given by explicit formulas but involving multidimensional Gaussian integrals. It is demonstrated that numerical integrations, using the methods available in MATLAB toolbox WAFO, give accurate approximations of the pdfs.

ACKNOWLEDGMENT

The first author was supported from Riksbankens Jubileumsfond Grant Dnr: P13-1024 and Swedish Research Council Grant Dnr: 2013-5180, while the second was partial support by Swedish Research Council Grant 340-2012-6004 and by Knut and Alice Wallenberg stiftelse.

REFERENCES

[1] S. Gran, *A Course in Ocean Engineering*, ser. Developments in Marine Technology. Amsterdam - London - New York - Tokyo: Elsevier Science Publishers, 1992, vol. 8.

[2] P. Brodtkorb, P. Johannesson, G. Lindgren, I. Rychlik, J. Ryden, and E. Sjö, "Wafo - a matlab toolbox for analysis of random waves and loads," in *Proceedings of the 10th International Offshore and Polar Engineering Conference*, vol. III, 2000, pp. 343–350.

[3] K. Podgórski, I. Rychlik, and U. Machado, "Exact distributions for apparent waves in irregular seas," *Ocean Eng.*, vol. 27, pp. 979–1016, 2000.

[4] J. Rydén, S. van Iseghem, M. Olagnon, and R. I., "Evaluating height-lenght joint distributions for the crest of ocean waves," *Applied Ocean Research*, vol. 24, pp. 189–201, 2002.

[5] M. Kac, "On the average number of real roots of a random algebraic equation." *Bull. Amer. Math. Soc.*, vol. 49, pp. 314–320, 1943.

[6] S. O. Rice, "The mathematical analysis of random noise," *Bell Syst Tech J*, vol. 23, no. 3, pp. 282–332, 1944.

[7] ——, "The mathematical analysis of random noise," *Bell Syst Tech J*, vol. 24, no. 1, pp. 46–156, 1945.

[8] M. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, 1983.

[9] I. Rychlik, "Note on cycle counts in irregular loads," *Fatigue Fract. Engng Mater. and Struc.*, vol. 16, pp. 377–390, 1993.

[10] ——, "A new definition of the rainflow cycle counting method," *Int. J. Fatigue*, vol. 9, pp. 119–121, 1987.

[11] R. Willson and I. Rychlik, "Segmentation of mineral textures," Image Segmentation Workshop 96, The Australian Pattern Recognition Society, Sydney Australia, can be downloaded from Research Gate, 1996.

[12] R. Willson and D. Nott, "Review of recent results on excursion set models," *Image Anal Stereol*, vol. 20, pp. 71–78, 2001.

[13] U. Zähle, "A general Rice formula, Palm measures, and horizontal-window conditioning for random fields," *Stochastic Processes and their Applications*, vol. 17, pp. 265–283, 1984.

[14] J.-M. Azaïs and M. Wschebor, *Level Sets and Extrema of Random Processes and Fields*. New York: Wiley, 2009.

[15] M. Longuet-Higgins, "The statistical analysis of a random moving surface," *Phil. Trans. Roy. Soc. London, Series A*, vol. 249, pp. 321–387, 1957.

[16] A. Baxevani, K. Podgórski, and I. Rychlik, "Velocities for moving random surfaces," *Probabilistic Engineering Mechanics*, vol. 18, pp. 251–271, 2003.

[17] R. Adler, *The Geometry of Random Fields*. New York: Wiley, 1981.

[18] A. Baxevani and I. Rychlik, "Maxima for gaussian seas," *Ocean Engineering*, vol. 33, pp. 895–911, 2006.

[19] C. Mercadier, "Numerical bounds for the distributions of the maxima of one- and two-parameter gaussian processes." *Adv. Appl. Prob.*, 2006.

[20] G. Lindgren and I. Rychlik, "Slepian models and regression approximations in crossing and extreme value theory," *International Statistical Review*, vol. 59, pp. 195–225, 1991.

[21] P. Brodtkorb, "Evaluating nearly singular multinormal expectations with application to wave distributions," *Methodology and Computing in Applied Probability*, vol. 8, pp. 65–91, 2006.

http://journals.lub.lu.se/stat

LUND UNIVERSITY
School of Economics and Management

Working Papers in Statistics 2016
LUND UNIVERSITY
SCHOOL OF ECONOMICS AND MANAGEMENT
Department of Statistics
Box 743
220 07 Lund, Sweden