Deriving environmental contours from highest density regions

Andreas F. Haselsteiner\textsuperscript{a,b,*}, Jan-Hendrik Ohlendorf\textsuperscript{a,b}, Werner Wosniok\textsuperscript{c}, Klaus-Dieter Thoben\textsuperscript{a,b,d}

\textsuperscript{a}University of Bremen, Faculty of Production Engineering, BIK – Institute for Integrated Product Development, 28359 Bremen, Germany
\textsuperscript{b}ForWind – Center for Wind Energy Research of the Universities of Oldenburg, Hannover and Bremen, 26129 Oldenburg, Germany
\textsuperscript{c}University of Bremen, Faculty of Math and Computer Science, Institute of Statistics, 28359 Bremen, Germany
\textsuperscript{d}BIBA - Bremer Institut für Produktion und Logistik GmbH, 28359 Bremen, Germany

Abstract

Environmental contours are an established method in probabilistic engineering design, especially in ocean engineering. The contours help engineers to select the environmental states which are appropriate for structural design calculations. Defining an environmental contour means enclosing a region in the variable space which corresponds to a certain return period. However, there are multiple definitions of environmental contours for a given return period as well as different methods to compute a contour. Here, we analyze the established approaches and present a new concept which we call highest density contour (HDC). We define this environmental contour to enclose the highest density region (HDR) of a given probability density. This region occupies the smallest possible volume in the variable space among all regions with the same included probability, which is advantageous for engineering design. We perform the calculations using a numerical grid to discretize the original variable space into a finite number of grid cells. Each cell’s probability is estimated and used for numerical integration. The proposed method can be applied to any number of dimensions, i.e. number of different variables in the joint probability model. To put the highest density contour method in context, we compare it to the established inverse first-order reliability method (IFORM) and show that for common probability distributions the two methods yield similarly shaped contours. In multimodal probability distributions, however, where IFORM leads to contours which are difficult to interpret, the presented method still generates clearly defined contours.

Keywords: highest density contour (HDC), joint probability distribution, numerical integration, sea state, engineering design, inverse first order reliability method (IFORM)

Nomenclature

| Symbol | Description |
|--------|-------------|
| $\alpha$ | Exceedance probability [-] |
| $\alpha_W, \beta_W, \gamma_W$ | Parameters of a Weibull distribution [-] |
| $\bar{f}$ | Cell-averaged joint probability density [-] |
| $\bar{F}(f_m)$ | Probability enclosed by a contour of $f_m$ probability density [-] |
| $\bar{f}_{X1}$ | Cell-averaged probability density in dimension 1 [-] |
| $\bar{f}_{X2|X1}$ | Cell-averaged probability density in dimension 2 conditional on $x_1$ [-] |
| $\beta$ | Radius in $U$-space used in IFORM [-] |
| $\mathcal{F}$ | Failure region [-] |
| $\mu_2, \sigma_2$ | Parameters of a normal distribution [-] |
| $\theta$ | Angle [deg] |
| $\tilde{\mu}_{Hs}, \tilde{\sigma}_{Hs}$ | Parameters of a log-normal distribution [-] |
| $a_1, a_2, a_3, b_1, b_2, b_3$ | Fitted parameters of the conditional distribution [-] |
| $C$ | Set making up the environmental contour [-] |
| $F()$ | Cumulative distribution function [-] |
| $f()$ | Probability density function [-] |
| $f_m$ | Minimum probability density of the enclosed region / constant probability density along the contour [-] |
| $f_m^*$ | Normalized minimum probability density [-] |
| $H_s$ | Significant wave height, random variable [m] |
| $h_s$ | Significant wave height, realization [m] |
| $H_{s,25}$ | 25-year return value of the significant wave height based on its marginal distribution [m] |
| $H_{s,\theta}$ | Maximum significant wave height along the contour [m] |
1. Introduction

1.1. Purpose of environmental contours

Engineers have to design any marine structure in such a way that it is able to withstand the loads induced by the environment. As the environment, i.e. wind, waves and currents, continually change and cannot be predicted for long periods of time, the environment is often modeled stochastically by defining probability density functions, \( f(x_j) \). Then, the structure is designed to withstand all but some extremely rare environmental states, e.g. all waves with significant wave heights, \( H_s \), less than a threshold, \( h_s \), with a cumulative probability or exceedance probability of \( \alpha \), i.e. \( Pr(H_s \leq h_s) = 1 - \alpha \) or \( Pr(H_s > h_s) = \alpha \). In general notation for any random variable, \( X_1 \), there exists a threshold, \( x_1 \), which fulfills

\[
F(x_1) = Pr(X_1 \leq x_1) = \int_{-\infty}^{x_1} f(x)dx = 1 - \alpha. \tag{1}
\]

The exceedance probability, \( \alpha \), corresponds to a certain recurrence or return period, \( T \), which describes the average time period between two consecutive environmental states above the threshold, \( x_1 \). The threshold is called return value. For example, to comply with standards a marine structure such as an offshore wind turbine is required to withstand significant wave heights, \( H_s \), with a return period, \( T \), of 50 years \([20]\).

Often, however, structural safety depends not only on one variable, but on the occurrence of combinations of \( p \) variables, \( \{X_j\}_{j=1}^{p} \). When two variables are of importance, e.g. significant wave height, \( H_s \), and spectral peak period, \( T_p \), a joint probability density function can be defined and an environmental contour can be calculated which encloses the subset (or region) of environmental states that the structure has to be designed for. Here, we call this region design region (Fig. 1). Often the most critical structural response is associated with very high or low values of environmental variables, i.e. with environmental conditions located at the boundary of the design region. Consequently, standards allow engineers to calculate structural responses for a limited set of environmental design conditions along the contour instead of requiring engineering calculations based on a high number of possible variable combinations spread over the complete design region \([8]\). If there are more than two variables the concept of environmental contours leads to environmental surfaces (3 variables) or environmental manifolds (> 3 variables). Here, for simplicity we also refer to these as environmental contours.

1.2. Different definitions and methods

As there are different mathematical definitions for environmental contours one has to further specify which kind of environmental contour is being calculated. Different concepts of environmental contours lead to different design loads and consequently to different structural responses.
Originally, environmental contours arose from the concept of return values in univariate probability density functions which are calculated based on one-sided exceedance over threshold (Fig. 2a). Consequently, a logical definition for an environmental contour is (i) constant one-sided exceedance in all directions of the p-dimensional variable space, \( Pr(X_1 > x_1, X_2 > x_2, ..., X_p > x_p) = \alpha \). The bottom panel in Fig. 2a shows the contour for the two-dimensional joint distribution of \( X_1 \) and \( X_2 \). However, for design purposes not only the highest values of a variable can be of interest, but also the lowest. For example, when designing an offshore structure, low values of the peak period, \( T_p \), have to be considered as the structure’s natural frequencies can be either higher or lower than the average peak period. Consequently, another possible definition for an environmental contour is (ii) two-sided exceedance over threshold (Fig. 2b; e.g. [21]). A third possibility is to define an environmental contour to have (iii) constant probability density, \( f_m \), along its path enclosing the most likely environmental states (Fig. 2c). In this case a \( T \)-year return period means that on average every \( T \) years an environmental state with a probability density less than \( f_m \) occurs. In the broader statistics literature the variable region enclosed by such a contour is called a highest density region (HDR) [19]. Although HDRs are a logical concept for environmental contours, yet no author has strictly followed this definition. The design curve introduced by Haver [14] is a related concept since it is a line of constant probability density, but only one-sided exceedance is considered. The constant probability density approach described by Det Norske Veritas [8] does define a fully closed contour of constant probability density. However, it is designed in such a way that it is unclear how much probability is enclosed by the contour. Instead the contour’s probability density, \( f \), is chosen to be the joint probability density of the \( (x_1, x_2) \)-variable combination with \( x_1 = \text{return value based on the marginal } x_1 \text{-distribution and } x_2 = \text{an associated } x_2 \text{ value (Fig. 3c). Leira [23], however, has indeed used the HDR definition but only after a transformation of the original variables into standard normal space. When transforming the contour back to the original variable space the constant probability density is not preserved. Here we will compute contours strictly following the HDR definition.

Besides these different definitions of types of environmental contours there exist different methods to calculate a given type of environmental contour. The traditional and probably most used approach is the so-called inverse first-order reliability method (IFORM) [34, 15]. It is a standard design practice for a wide range of marine engineering applications where extreme sea states are of interest [8]. These are for example ships [11], offshore wind turbines [20], floating structures [9] or wave energy converters [7, 16]. Using IFORM one defines the contour in standard normal space, \( U_j \), instead of the original environmental variable space, \( X_j \). Thus, one first defines a circle with a radius, \( \beta \), in the \( U \)-space (Fig. 3a). The radius corresponds to the return period and increases with longer periods. Then one transforms the points along the circle to the original variable space leading to the environmental contour. This transformation is done via the inverse Rosenblatt transformation [27]. As its name implies IFORM is a reliability method and is based on the idea that the exceedance region approximates the failure region, \( F \), of a structure (and the exceedance probability, \( \alpha \), approximates the structure’s failure probability, \( P_f \); see [25]). Contours based on IFORM are widely used and have been published e.g. by Saranyasoontorn and Manuel [28], Leira [23], Baarholm et al. [2], Li et al. [24], Myers et al. [26], Valamanesh et al. [30], Eckert-Gallup et al. [13].

Huseby et al. [17], however, pointed out that the Rosenblatt transformation introduces errors as failure probabilities, \( P_f \), can be underestimated or overestimated on a case by case basis. Therefore, they introduced an alternative
Established methods to calculate environmental contours. (a) Inverse first-order reliability method (IFORM) [34]. The contour is defined as a circle in standard normal variables, $U_j$. The points along the circle have to be transformed to the original environmental variables, $X_j$. (b) Huseby et al. [17] Monte Carlo contour. The contour is computed with the environmental variables, $X_j$, directly. (c) Constant probability density contour described by Det Norske Veritas [8]. By its definition it is unclear how much probability is enclosed by the contour. (d) Jonathan et al. [21] constant exceedance probability contour. The calculation can be done in a general set of variables, $M_j$, e.g. in the $X$- or $U$-space. In comparison to the other methods a different definition for the exceedance region is used (compare shaded areas).

Method to calculate environmental contours in the original variable space. Following their method, one first carries out a Monte Carlo simulation to generate a high number of sea states based on a given joint probability distribution model. Then one chooses an angle, $\theta$, defining a line (in two dimensions, $p=2$) and varies its position such that it divides the variable space into one halfspace containing the majority of data points and the other halfspace containing the data points representing the exceedance probability, $\alpha \times n$ (with $n$ being the total number of simulated environmental states, Fig. 3b). By iterating this procedure over a finite number of angles, $\theta \in [0,360)$, the resulting lines can be connected to an environmental contour. This new approach has been picked up in several recent publications, e.g. to compare the approach to the traditional IFORM method [33], to compare different statistical models [31] or to decrease the required process time [18]. While the Monte Carlo method overcomes the problems caused by the Rosenblatt transformation it requires the simulation of environmental states which is computationally more expensive than the simple IFORM calculations. Further, by its definition the method cannot generate concave contours.

Jonathan et al. [21] define and calculate environmental contours yet differently. Using clear mathematical notation they find a contour with constant exceedance probability, $\Pr(X_1 > x_1, X_2 > x_2) = \alpha$ (notation for two dimensions, $p=2$). Thus, instead of finding halfspaces which are tangential to the contour, their exceedance regions have finite boundaries for each variable leading to outwards radiating rectangles in a two-dimensional Cartesian coordinate system (Fig. 3d). Consequently, in contrast to IFORM and the Monte Carlo approach the method does not try to match the exceedance region with the failure region and thus separates the concept of an environmental contour from a structure’s failure function. Following this method one first chooses a reference point, $r_{M0}$. Then one defines a line which passes through that point at an angle, $\theta$, to the abscissa. Lastly one finds the position along the line which satisfies $Pr = \alpha$. Repeating this procedure over a full circle, $\theta \in [0,360)$, one finds the environmental contour. The method can be applied in any variable space, $M_j$, e.g. in original variables, $X_j$, or standard normal variables, $U_j$. Further, besides fully closed contours, one-sided exceedance is also considered by the authors. One can combine the method with using modern conditional extreme models [16] as demonstrated by Jonathan et al. [22, 21]. The method disconnects the environmental state statistics from any particular structural problem which makes it a more general approach to define a $T$-year set of environmental states for any further use of these data. However, like the reliability methods, it defines multiple $\alpha$-exceedance regions in the variable space of a single probability model. While in reliability methods the idea is that one of these multiple exceedance regions overlaps with the failure region this is not the case with the $\Pr(X_1 > x_1, X_2 > x_2) = \alpha$ definition. Thus, if a contour is defined independently of the concept of failure regions, it seems more meaningful to define $\alpha$ to be the probability of a single region (instead of having multiple regions with $\alpha$ probability content).

Motivated by the individual advantages the described contour calculation methods have, here, we introduce contours enclosing highest probability density regions which we compute using numerical integration. We continue the idea introduced by Jonathan et al. [21] of decoupling the exceedance region from the structure’s failure region, but go one step further and do not define any kind of outwards radiating exceedance region. Instead, we choose to find a contour which encloses the most likely environmental states which together make up a defined probability of $1 - \alpha$. The proposed method allows us to define the contour in the original variable space and can be used for any number $p$ of dimensions. By discretizing the variable space into a finite number of grid cells and using numerical integration techniques any probability distribution can be evaluated, e.g common parametric sea state joint probability distributions [32], nonparametric models [12] or extreme value models which can have discontinuities at the threshold [29]. Similar as being done e.g. in computational fluid dynamics [5] we demonstrate that with a sufficiently small grid cell size the solution is grid independent.
2. Data

In order to compare our environmental contour approach to similar methods we use the 3-hour sea state model presented by Vanem and Bitner-Gregersen [32]. They use a fitted joint model for significant wave height, $H_s$, and zero-upcrossing period, $T_z$. Based on their model environmental contours have been calculated using both the traditional IFORM method [32] and the newer Monte Carlo method [17]. The joint model was derived from one particular location in the ERA-Interim data set [6]. Significant wave height, $H_s$, is modeled as a 3-parameter Weibull distribution with the parameters $\alpha_W$ (scale), $\beta_W$ (shape) and $\gamma_W$ (location):

$$f_{H_s}(h_s) = \frac{\beta_W}{\alpha_W} \left( \frac{h_s - \gamma_W}{\alpha_W} \right)^{\beta_W-1} \exp \left[ - \left( \frac{h_s - \gamma_W}{\alpha_W} \right)^{\beta_W} \right ]; \quad h_s \geq \gamma_W. \quad (2)$$

Based on a least squares fit the parameters are $\alpha_W = 2.776$, $\beta_W = 1.471$ and $\gamma_W = 0.8888$ [32].

The zero-upcrossing period, $T_z$, is modeled to follow a log-normal distribution, LN:

$$f_{T_z|H_s}(t_z|h_s) = \frac{1}{t_z \sigma_{H_s} \sqrt{2\pi}} \exp \left[ \frac{-(\ln t_z - \bar{\mu}_{H_s})^2}{2\bar{\sigma}_{H_s}^2} \right]. \quad (3)$$

The distribution’s parameters, $\bar{\mu}_{H_s}$ and $\bar{\sigma}_{H_s}$, are conditional on the significant wave height, $H_s$, and are modeled as 3-parameter functions:

$$\bar{\mu}_{H_s}(h_s) = a_1 + a_2 h_s^{a_3}, \quad (4)$$

$$\bar{\sigma}_{H_s}(h_s) = b_1 + b_2 \exp(b_3 h_s). \quad (5)$$

In this case they are estimated to be $a_1 = 0.1000$, $a_2 = 1.489$, $a_3 = 0.1901$, $b_1 = 0.0400$, $b_2 = 0.1748$, $b_3 = -0.2243$ [32].

Multiplying the marginal distribution of the significant wave height, $H_s$, and the conditional distribution of the zero-upcrossing period, $T_z$, we can calculate the joint distribution:

$$f_{H_s,T_z}(h_s,t_z) = f_{H_s}(h_s)f_{T_z|H_s}(t_z|h_s). \quad (6)$$

Since the data represent 3-hour sea states, exceedance probability, $\alpha$, for a $T$-year return period is calculated as

$$\alpha = \frac{1}{n} = \frac{1}{T \times 365.25 \times 24/3}. \quad (7)$$

3. Highest density contour (HDC)

3.1. Analytical definition

Our goal is to find a contour, $C$, of constant probability density, $f_m$, which encloses a probability of $1 - \alpha$, i.e.:

$$C(f_m) = \{ x : x \in \mathbb{R}^p, f(x) = f_m \},$$

$$R(f_m) = \{ f(x) \geq f_m \},$$

$$\int_{R(f_m)} f(x)dx = 1 - \alpha. \quad (8)$$

This contour, $C$, encloses the highest density region, $R$. Therefore we call $C$ highest density contour (HDC). A highest density region fulfills two main properties: (i) the probability density of every point inside is at least as large as the probability density of any point outside and (ii) for a given probability content the region occupies the smallest possible volume in the variable space [4]. There is no general analytic solution to find the HDR or HDC, i.e. solving for $C$ or $R$ in Eq. 8.

HDRs, however, can be computed based on numerical integration approaches [35] or Monte Carlo techniques [19]. Environmental contours involve very low $\alpha$ values and are usually based on low-dimensional probability models. Thus, we choose numerical integration over Monte Carlo simulation to compute the highest density contour, $C$. However, if a probability model, which incorporates many environmental variables (high $p$ value), is evaluated numerical integration might become infeasible and Monte Carlo approaches should be used. Here, we use numerical integration and start by discretizing the probability density space into a finite number of equally sized grid cells. In the next section we will evaluate the two-dimensional case, but in the appendix the equations for $p$ dimensions are given.

3.2. Numerical integration approach in two dimensions

The two-dimensional probability space is discretized in $K \times L$ grid cells with a constant size of $\Delta x_1 \times \Delta x_2$ (Fig. 4). Each grid cell’s center point, $(x_{1k}, x_{2l})$, is used as the reference position of the cell. Then, based on the cumulative distribution function, $F_{X_1}$, the cell-averaged probability density in the first dimension, $\bar{f}_{X_1}$, is calculated using central difference:

$$\bar{f}_{X_1}(x_1) = \frac{F_{X_1}(x_1 + 0.5\Delta x_1) - F_{X_1}(x_1 - 0.5\Delta x_1)}{\Delta x_1}. \quad (9)$$

The cell-averaged probability density in the second dimension, $\bar{f}_{X_2|X_1}$, is calculated similarly:

$$\bar{f}_{X_2|X_1}(x_2|x_1) = \frac{F_{X_2|X_1}(x_2 + 0.5\Delta x_2) - F_{X_2|X_1}(x_2 - 0.5\Delta x_2)}{\Delta x_2}. \quad (10)$$

While $\bar{f}_{X_1}$ is the true cell-averaged probability density in the first dimension, in the second dimension, $\bar{f}_{X_2}$ is
approximated since the dependence of $F_{x_2|x_1}$ upon $x_1$ within the grid cell is not accounted for. Instead we fix $x_1$ to the value at the grid cell center, $x_1 = x_{1l}$, and therefore assume $F_{x_2|x_1}$ to be constant from $x_{1l} - 0.5 \Delta x_1$ to $x_{1l} + 0.5 \Delta x_1$.

Multiplying the two individual probability densities yields the cell-averaged joint probability density, $\bar{f}$:

$$\bar{f}(x_1, x_2) = \bar{f}(x_1) \bar{f}_{x_2|x_1}(x_2|x_1).$$

Now we can compute the probability that an event with a minimum probability density of $f_m$ occurs, i.e. we calculate the probability content enclosed by a HDC of $f_m$ probability density. This probability, $\bar{F}(f_m)$, is calculated by summing up the probabilities of all cells which have a probability density greater than or equal to $f_m$ (Fig. 4):

$$\bar{F}(f_m) = \sum_{k=1}^{K} \sum_{l=1}^{L} \left\{ \begin{array}{ll} \bar{f}(x_{1k}, x_{2l}) \Delta x_1 \Delta x_2 & \text{if } \bar{f}(x_{1k}, x_{2l}) \geq f_m \\ 0 & \text{if } \bar{f}(x_{1k}, x_{2l}) < f_m. \end{array} \right.$$  \hfill (12)

If the joint probability density function is unimodal the grid cells which fulfill $\bar{f} \geq f_m$ make up a single contiguous area. The boundary of this area is a contour which encloses a probability of $\bar{F}$. Using the function $\bar{F}(f_m)$ we can consequently find a contour with a given exceedance probability, $\alpha$, of interest by finding the corresponding minimum probability density, $f_m$:

$$\bar{F}(f_m) = 1 - \alpha.$$  \hfill (13)

Solving this equation is a root finding problem of a monotonically decreasing function ($\bar{F}(f_m) = 1 + \alpha = 0$). We solve the equation using Matlab’s (version R2015b, The MathWorks, USA) fzero function which iteratively finds the root of a nonlinear function. All grid cells fulfilling $\bar{f} \geq f_m$ then approximate the HDR, $R(f_m)$, and the grid cells making up the boundary of the HDR approximate the HDC, $C(f_m)$.

4. Results and discussion

4.1. Properties of the highest density contours

As done in previous work based on the described joint probability model [17, 32] we compute the 1-, 10- and 25-year environmental contours (Fig. 5). The corresponding exceedance probabilities are $\alpha_1 = 3.42 \times 10^{-4}$, $\alpha_{10} = 3.42 \times 10^{-5}$ and $\alpha_{25} = 1.37 \times 10^{-5}$ respectively. The computed HDCs have constant probability densities of $f_{m1} = 4.4 \times 10^{-5}$ (1-year), $f_{m10} = 4.3 \times 10^{-6}$ (10-year) and $f_{m25} = 1.7 \times 10^{-6}$ (25-year). Fig. 6a shows how the enclosed probability, $\bar{F}$, monotonically decreases with increasing $f_m$ until it reaches $\bar{F} = 0$. Since the probability functions we use here (Weibull and log-normal) are unbounded, $\bar{F}$ asymptotically approaches 1 as $f_m$ approaches 0. Fig. 6b presents the maximum $H_s$- and $T_s$-values along a contour of constant $f_m$-probability density ($H_{s \leq}, T_{s \geq}$). Longer return periods, $T$, lead to smaller $f_m$-values and consequently to bigger contours with higher $H_{s \leq}$ and $T_{s \geq}$ values.

As discretization in general is sensitive to step size we evaluate the contour’s robustness with respect to grid cell size $\Delta x_1 = \Delta H_s$, $\Delta x_2 = \Delta T_s$. We analyze how minimum probability density, $f_m$, changes with grid cell size. In all three tested return periods (1-, 10- and 25-year contour) minimum probability density, $f_m$, is roughly constant at small cell sizes and starts to fluctuate with increasing cell size indicating a grid-independent solution can be reached (Fig. 7a). Oversized grid cells can lead to minimum probability density being half or double than the converged minimum probability density (Fig. 7b). For the given probability model we find that convergence is reached at a grid cell size of $H_s = 0.05 m$ and $T_s = 0.05 s$. There, deviation to the smallest tested grid cell size is less than 1%, 0.99 < $f_m$ < 1.01, with $f_m$ being minimum probability density, $f_m$, normalized by the converged $f_m$ value (Fig. 7c).

4.2. Comparison with IFORM and Monte Carlo contours

For comparison we further compute environmental contours using IFORM based on the same probability model. The highest density contours have similar shapes as the contours calculated with IFORM and the Monte Carlo method (Fig. 8c,d). However, we define a HDC to enclose a probability of $1 - \alpha$ while an IFORM contour and a Huseby et al. [17] Monte Carlo contour each enclose a probability less than $1 - \alpha$ since by their definitions multiple regions outside the contour have a probability of $\alpha$ (Fig. 8a). Consequently, the HDC’s dimensions in terms of $H_s$ and $T_s$ are bigger in comparison. However, for a fairer comparison we can inflate an IFORM contour and find
the $T$-year contour which encloses exactly $1 - \alpha$ probability. Leira [23] showed that this can be done by utilizing the inverted Rayleigh distribution (for two dimensions). The author calls these contours *equi-shape contours*. Here, we find that such a 25-year equi-shape contour corresponds to a 308.8-year IFORM contour. The contour’s shape and size is roughly similar to the 25-year HDC. These similarities suggest that the 308.8-year IFORM contour has approximately constant probability density, $f_m^{25}$, along the contour.

To visualize a typical data set, we Monte Carlo simulate 25 years of 3-hour sea states ($\alpha = 73050$; gray dots in Fig. 8a). In this particular data set one data point exceeds the HDC while there are multiple data points exceeding the 25-year IFORM contour. The different contour dimensions can also be expressed in terms of maximum $H_s$- and $T_z$-values along the contour ($H_s^{\dot{f}}$, $T_z^{\dot{f}}$). While Huseby et al. [17] report 25-year maxima of $H_s^{25} = 14.66$ m and $T_z^{25} = 13.68$ s for the Monte Carlo contour, here we find $H_s^{\dot{f}} = 16.79$ m and $T_z^{\dot{f}} = 14.64$ s for the HDC and $H_s^{\dot{f}} = 15.23$ m and $T_z^{\dot{f}} = 13.96$ s for the IFORM contour (Fig. 8d). Thus, the HDC $H_s^{\dot{f}}$ value is 10.2 % higher than the IFORM value and 14.5 % higher than the Monte Carlo method value. Consequently, from an engineering design point of view the HDC is the most conservative method of the three considered.

This does not only apply to the considered distribution, but is a generic property based on the different definitions of these three contours. The IFORM and Monte Carlo contours are defined to contain the return value of the marginal distribution as their highest variable value, i.e. $H_s^{x,25} = H_{s,25}$ (Fig. 8a). On the other hand, a HDC is defined to enclose $1 - \alpha$ probability. Since it does not contain all $H_s-T_z$ sea states fulfilling $H_s < H_{s,25}$ (which together would make up $1 - \alpha$ probability) it must contain some sea states with $H_s > H_{s,25}$.

By the HDC’s definition of an enclosed probability of $1 - \alpha$, in a random 25-year data set the probability that at least one data point exceeds a 25-year contour is about 63.2 %, $1-(1-\alpha_{25})^n \approx 0.632$ with $n = 25 \times 365.25 \times 24/3 = 73050$. Here, exceedance precisely means that this sea state realization is anywhere outside the region enclosed by the contour, $R(f_m)$. Such a sea state occurs on average every 25 years. This simple and clear interpretation is why we have chosen the definition of constant probability density and a probability of $1 - \alpha$, i.e. defining the contour to enclose the highest density region. We believe that this definition offers an intuitive and meaningful concept for a $T$-year environmental contour in the engineering design process. If an engineer designs a structure to withstand all sea states inside a $T$-year contour, the structure will be designed for the most likely (extreme) sea states which are expected to occur in $T$ years. Then on average every $T$ years a sea state will occur which the structure is not designed for.

Alternative concepts with multiple $\alpha$-exceedance regions (see Fig. 3a, b) are based on the idea of known failure regions in the context of structural reliability methods (see [25]). IFORM assumes that a structure’s *failure surface* (or *limit state surface*) has a convex shape. It defines the $\alpha$-halfspace exceedance regions in its particular way because in that case the true failure surface can be linearized such that the variable space is separated by a straight line at an angle $\theta$ into a survival region and a failure region (in two dimensions). Then, this failure region overlaps with IFORM’s exceedance region. It has the failure probability $P_f = \alpha$ and the survival region the survival probability $1 - P_f$. Here, however, we completely separate the idea of describing the environmental conditions from any particular structural problem. Thus, we do not intend to align the $\alpha$-probability exceedance region with a particular failure region.

As described IFORM leads to a contour which encloses less than $1 - \alpha$ probability and consequently results in less conservative design conditions compared to a HDC. If the structural design, which is developed based on these environmental conditions, has a convex failure surface, the theoretical precondition of IFORM is met. Then in comparison, a HDC can be seen as overly conservative. Thus, if the designer knows that a structure responds with a convex failure surface choosing an IFORM contour is advantageous in the sense that it yields less conservative but still safe design conditions.

While many structures respond with a convex failure surface this precondition for IFORM connects the environmental contour to a certain class of structures. The shape of the failure surface might be unknown beforehand and
Figure 6: Expansion of the highest density contour. (a) The probability enclosed by the contour, \( F(f_m) \), is 1 at a minimum probability density of \( f_m = 0 \) and monotonically decreases to \( F(f_m \approx 0.12) = 0 \). Probabilities corresponding to the 1-, 10- and 25-year contour are shown. The inlet illustrates the definition of \( F \) and \( f_m \). (b) Maximum variable values along the contour, \( H_{s\epsilon} \) and \( T_{z\epsilon} \), as a function of minimum probability density, \( f_m \). The inlet illustrates that there is no \( (H_{s\epsilon},T_{z\epsilon}) \)-sea state along the contour. Instead, the \( (H_{s\epsilon},T_{z\epsilon}) \)-sea state has a \( T_z \) value different from \( T_{z\epsilon} \) and vice versa.

Figure 7: Grid independence study. Quadratic grid cells with sizes ranging from 0.01 to 10 units of grid cell length are tested to evaluate grid convergence. (a) The contour’s minimum probability density, \( f_m \), for a given return period is sensitive to grid cell size \( \Delta H_s, \Delta T_z \). Sensitivity increases with grid cell length. (b) If grid cell size is too big minimum probability density, \( f_m \), can be half or double than the converged minimum probability density. Plotted is \( f_m^* \) which is minimum probability density, \( f_m \), normalized by the converged \( f_m \) value. (c) Aiming for grid convergence with an error of less than 1% we use grid cells with dimensions of 0.05 m×0.05 s (marked with a vertical line).
Figure 8: **Comparison of environmental contours derived using different methods.** (a) Sketches showing expected differences in contour size due to different definitions. Some contours are defined in such a way that the maximum value along the contour, $H_{s,25}$, is equal the return value of the marginal distribution, $H_{s,25}$, (middle). The highest density contour (HDC), however, is defined to enclose $1-\alpha$ probability and thus has a maximum value along the contour which is higher than the return value of the marginal distribution, $H_{s,25} > H_{s,25}$, (right). (b) Sketch illustrating an IFORM contour and possible failure regions of a linear system of three components, $F_1$, $F_2$, $F_3$. Since the contour contains less than $1-\alpha$ probability the system’s failure probability can be greater than $\alpha$. (c) A total of $n = 73050$ sea state data points have been Monte Carlo simulated representing a 25-year data set (scatter plot). The 25-year HDC (solid line) and the 25-year IFORM contour (short dashes) have similar shapes, but as expected the HDC is bigger. The 308.8-year IFORM contour or 25-year equi-shape contour (long dashes; [23]) encloses the same amount of probability as the 25-year HDC. (d) Comparison of maximum values along the contour, $H_{s,f}$ and $T_{z,f}$. As expected by the different definitions, the HDC has the highest maximum significant wave height, $H_{s,f}$, and maximum zero-upcrossing period, $T_{z,f}$. 
only becomes apparent during the design process. If it turns out that the failure surface is non-convex and therefore violates IFORM’s precondition the designer would need to go one step backwards and define new design conditions by inflating the IFORM contour. By not making use of the properties of possible structural responses the HDC is more conservative, but also more general in its application. It would avoid the need of the described iteration loop in the design process.

Further, a highest density contour is advantageous in the design process of a structural problem of a system consisting of multiple components. Consider a series structure consisting of \( z \) different components with \( z \) different failure functions. In a series structure a failure of one component results in failure of the system \([3]\). Suppose that each component fulfills IFORM’s precondition of having a convex failure surface. Nevertheless, the probability contained by the union of all \( z \) failure regions, \( F_1 \cup F_2 \cup \ldots \cup F_z \), could exceed \( \alpha \) (Fig. 8b). In that case it would be expected that fre quenter than every \( T \) years an environmental state occurs which leads to failure of some of the components and consequently failure of the system. If an environmental contour containing \( 1 - \alpha \) probability were used to design the components, on the other hand, by definition the system’s probability of failure would be less than \( \alpha \). Consequently, the system would be expected to survive longer than \( T \) years.

A similar example could be given for a single component with multiple failure modes. The three failure regions shown in Fig. 8b would then correspond to different failure modes and the same conclusions as for the series structure could be drawn. These two examples explain why IFORM is primarily aimed at assessing the reliability of one component failing in one particular failure mode. A highest density contour, on the other hand, could be used in these two cases without worrying that any assumptions might be violated.

4.3. Bi modal mixture model

Highest density contours can be computed based on any probability distribution. The used definition of constant probability density along the contour, \( f_m \), can lead to multiple enclosed subregions for a given return period, \( T \), if the probability distribution is multimodal. Here, we demonstrate this by extending the joint \( H_s-T_z \)-probability distribution by a mixture model for the zero-upcrossing period, \( T_z \). We use the \( H_s-T_z \) environmental variables although we are aware that such a \( H_s-T_z \) distribution might be physically unrealistic. However, for simplicity we build upon the previously described sea state model instead of setting up a new case with a different set of environmental variables. Thus, we keep the log-normal distribution term, \( LN(\bar{\mu}_{Hs}, \bar{\sigma}_{Hs}^2) \), from Eqs. 3–5 and mix it with a normal distribution, \( N(\mu_2, \sigma_2^2) \):

\[
f_{T_z|H_s}(t_z|h_s) = p_{H_s}LN(\bar{\mu}_{Hs}, \bar{\sigma}_{Hs}^2) + (1 - p_{H_s})N(\mu_2, \sigma_2^2).
\]

Similar to the parameters \( \bar{\mu}_{Hs} \) and \( \bar{\sigma}_{Hs} \), we define the mixture coefficient, \( p_{H_s} \), to be conditional on significant wave height, \( H_s \). Using an exponential decay function, we let the normal distribution term, \( N(\mu_2, \sigma_2^2) \), fade out at high significant wave height, \( H_s \):

\[
p_{H_s}(h_s) = 1 - \exp(-3h_s).
\]

We design two mixture models. For the first model we create a normal distribution, \( N \), such that its probability density blends smoothly into the log-normal distribution, \( LN \), by using a mean value of \( \mu_2 = 10 \) s and standard deviation of \( \sigma_2 = 2 \) s (model 1). For the second model we design a normal distribution which has much less density overlap by using a mean value of \( \mu_2 = 15 \) s and standard deviation of \( \sigma_2 = 0.5 \) s (model 2). For both models we compute the 25-year HDC as well as the 25-year IFORM contour. In model 1 the HDC and IFORM contour have similar shapes. Both have a concave path at high \( T_z \)-values and as expected the HDC is bigger in size (Fig. 9a). In contrast, model 2 has two distinct probability density maxima which lead to different shapes for the IFORM and HDC. While the HDC encloses two separated subregions the IFORM contour encloses a single contiguous region (Fig. 9b). This single region contains sea states with much lower probability densities than the conservative HDC as by its definition IFORM can only enclose one single contiguous region. Consequently, in this example an engineer who designs a structure to withstand all loads caused by the environmental states along this 25-year contour would design the structure to withstand some environmental states which are expected to occur extremely rarely. Therefore, possible structural designs which are limited by these environmental states would not be considered which could lead to bad design, either from a cost or engineering perspective.

The apparent difference in shape between the two contours is interesting since it visually demonstrates that the IFORM contour does not have constant probability along its path and consequently does not enclose the most likely environmental states. Strictly, this should not be expected anyway, but since it is roughly true for many ordinary sea state models, one might intuitively interpret an IFORM contour that way. By IFORM’s definition the contour has two properties in the \( U \)-space: (i) constant probability density along its path and (ii) \( \alpha \)-probability halfspaces separated by lines which are tangent to the contour (Fig. 3a). Interestingly, for many sea state probability models these two properties roughly translate to the \( X \)-space. Here, we demonstrate the robust persistence of the constant probability density property for the unmodified sea state model since in this case IFORM and HDCs have similar shapes (Fig. 8b). Rough persistence of the \( \alpha \)-halfspace property, on the other hand, has been shown by Huseby et al. \( [17] \) who computed Monte Carlo contours which are defined by enforcing the \( \alpha \)-halfspace property in the original variable space (Fig. 3b). These Monte Carlo contours have been
Figure 9: Environmental contours for mixture models. (a) Model 1 has a normal $T_z$-distribution, $N(\mu_2 = 10 \, \text{s}, \sigma_2 = 2 \, \text{s})$, which smoothly blends into the original $T_z$-log-normal distribution. The highest density contours and IFORM contours have similar shapes. (b) Model 2 has a normal $T_z$-distribution, $N(\mu_2 = 15 \, \text{s}, \sigma_2 = 0.5 \, \text{s})$, which leads to a second probability density maximum. Consequently, the highest density contour encloses two separated subregions. Due to its definition IFORM, however, encloses a single contiguous region.

reported to have similar shapes as the IFORM contours. Thus, based on experience an engineer might intuitively interpret a typical IFORM contour to have roughly constant probability density and $\alpha$-halfspace exceedance probability in the original variable space.

This interpretation would not hold true for the multimodal model 2, however. In addition to clearly not having constant probability density it also does not roughly have $\alpha$-halfspace exceedance in the original variable space since the contour is concave. Not having any meaningful properties in the original variable space, raises the question how to intuitively interpret an IFORM contour in such a case. In contrast, the presented highest density contour with its constant probability density, $f_m$, along the contour and its enclosure of a probability of $1 - \alpha$ offers a clear interpretation for any probability distribution.

5. Summary and conclusions

In this work we present environmental contours which enclose regions of highest probability density. A highest density contour (HDC) has constant probability density along its path and occupies the smallest possible volume in the variable space for a given probability content. We compute the contour using numerical integration based on a grid, i.e. we iteratively find the minimum probability density, $f_m$, along the contour and its enclosure of a probability of $1 - \alpha$. Defined this way a $T$-year environmental contour is exceeded on average every $T$ years anywhere along the contour. This means precisely that such an environmental state is realized anywhere outside the environmental contour (and not in a further limited exceedance region). Highest density contours can be computed based on any probability density function, e.g. standard parametric sea state models, nonparametric models or extreme value models. The method’s clear definition in terms of exceedance probability, $\alpha$, as well as its straightforward computation makes it an attractive alternative to the established IFORM approach.

Acknowledgements

We thank R. Reisenhofer for fruitful discussions. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Armstrong, C., Chin, C., Penesis, I., Drobyshevski, Y., 2015. Sensitivity of vessel responses to environmental contours of extreme sea states. In: ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers, p. V003T02A051.

[2] Baarholm, G. S., Haver, S., Øakland, O. D., 2010. Combining contours of significant wave height and peak period with platform response distributions for predicting design response. Marine Structures 23, 147–163.
wave spectral parameters for extreme sea states. Ocean Engineering 37, 1070–1080.

[23] Leira, B. J., 2008. A comparison of stochastic process models for definition of design contours. Structural Safety 30, 493–505.

[24] Li, L., Gao, Z., Moan, T., 2015. Joint distribution of environmental condition at five European offshore sites for design of combined wind and wave energy devices. Journal of Offshore Mechanics and Arctic Engineering 137, 031901.

[25] Madsen, H. O., Krenk, S., Lind, N. C., 1986. Methods of structural safety. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

[26] Myers, A. T., Arwade, S. R., Valamanesh, V., Hallof, S., Carswell, W., 2015. Strength, stiffness, resilience and the design of offshore wind turbine monopiles. Engineering Structures 100, 332–341.

[27] Rosenblatt, M., 1952. Remarks on a multivariate transformation. The Annals of Mathematical Statistics 23 (3), 470–472.

[28] Sarangasanthorn, K., Manuel, L., 2004. Efficient models for wind turbine extreme loads using inverse reliability. Journal of Wind Engineering and Industrial Aerodynamics 92, 789–804.

[29] Scarrott, C., MacDonald, A., 2012. A review of extreme value threshold estimation and uncertainty quantification. REVSTAT – Statistical Journal 10 (1), 33–40.

[30] Valamanesh, V., Myers, A. T., Arwade, S. R., 2015. Multivariate analysis of extreme metocean conditions for offshore wind turbines. Structural Safety 55, 60–69.

[31] Vanem, E., 2016. Joint statistical models for significant wave height and wave period in a changing climate. Marine Structures 49, 180–205.

[32] Vanem, E., Bitner-Gregersen, E. M., 2012. Stochastic modelling of long-term trends in the wave climate and its potential impact on ship structural loads. Applied Ocean Research 37, 235–248.

[33] Vanem, E., Bitner-Gregersen, E. M., 2015. Alternative environmental contours for marine structural design – a comparison study. Journal of Offshore Mechanics and Arctic Engineering 137 (5), 051601.

[34] Winterstein, S., Ude, T., Cornell, C., Bjoraker, P., Haver, S., 1993. Environmental parameters for extreme response: inverse FORM with omission factors. In: Proceedings, ICOSSAR-93. Innsbruck, Austria.

[35] Wright, D. E., 1986. A note on the construction of highest posterior density intervals. Applied Statistics 35 (1), 49–53.

Appendix A. Equations for p dimensions

We start by discretizing the p-dimensional probability space into $I_{X_j}^{p} K_j$ grid cells with grid cell lengths of $\Delta x_j$. Next, we calculate the cell-averaged probability density in each dimension, $f_{X_j}$. This is done based on the respective cumulative distribution function, $F_{X_j}$:

$$ f_{X_j}(x_j) = \frac{F_{X_j}(x_j + 0.5\Delta x_j) - F_{X_j}(x_j - 0.5\Delta x_j)}{\Delta x_j} $$

(A.1)

Multiplying the p individual probability densities yields the cell-averaged joint probability density, $f$:

$$ f(x_1, x_2, ..., x_p) = \prod_{j=1}^{p} f_{X_j}(x_j). $$

(A.2)

Next, we compute the probability enclosed by a contour of $f_m$ probability density. This is done by calculating the sum of each cell’s probability whose probability density is
greater than or equal $f_m$: 

$$F(f_m) = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \ldots \sum_{k_p=1}^{K_p} \begin{cases} \bar{f}(x_{1k_1}, x_{2k_2}, \ldots, x_{pk_p}) \Pi_{j=1}^{p} \Delta x_j & \bar{f} \geq f_m \\ 0 & \bar{f} < f_m \end{cases} \tag{A.3}$$

Now, we can proceed as in two dimensions. We want to find the minimum probability density, $f_m$, that corresponds to the exceedance probability, $\alpha$, of interest:

$$\bar{F}(f_m) = 1 - \alpha \tag{A.4}$$

As in two dimensions, this equation represents a root finding problem of a monotonically decreasing function ($\bar{F}(f_m) - 1 + \alpha = 0$) which can be solved with standard numerical methods, e.g. by using Matlab’s `fzero` function.

A Matlab implementation working up to four dimensions can be downloaded at http://mathworks.com/matlabcentral/fileexchange/60876. Figure A.10 shows a source code snippet and the corresponding flowchart.
fbarjoint = jointCellAveragedPdf(ModelArray, gridCenterPoints);

Fbarzero = @(fm)probabilityOfHdr(fbarjoint, fm, cellSize) - 1 + alpha;
fm = fzero(Fbarzero, 0);

hdrBinary = fbarjoint >= fm;

[x1, x2, x3, x4] = computeHdrBoundary(hdrBinary, gridCenterPoints);

Figure A.10: Computer program to derive a highest density contour. Left: Code snippet written in the Matlab programming language. Right: Corresponding flowchart.