Environmental contours

Arne Bang Huseby and Kristina Rognlien Dahl

University of Oslo, Norway

STK 4400

Environmental contours

- Let $(T, H) \in \mathbb{R}^2$ be a vector of environmental variables where e.g.,:

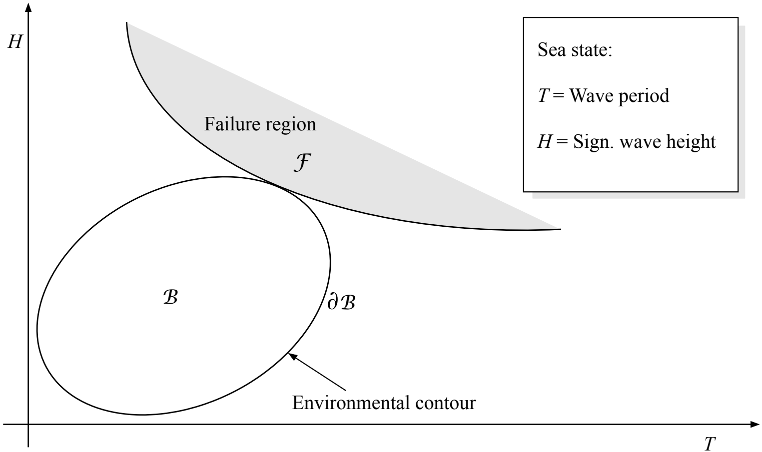
$$\begin{aligned} T &= \text{Wave period} \\ H &= \text{Significant wave height} \end{aligned}$$

- The distribution of (T, H) is assumed to be absolutely continuous with respect to the Lebesgues measure in $\mathbb{R}^2$.
- An *environmental contour* is then defined as the boundary of a set $\mathcal{B} \subseteq \mathbb{R}^2$, and denoted $\partial\mathcal{B}$.

Contours and failure regions

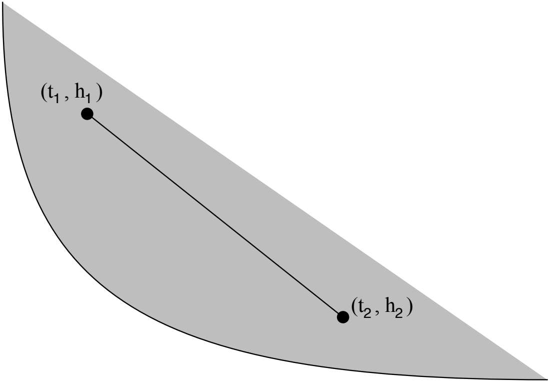
- The environmental contour $\partial\mathcal{B}$ represents possible design requirements for some structure of interest.
- The *failure region* $\mathcal{F} \subseteq \mathbb{R}^2$ of a structure is the set of states of the environmental variables where the structure fails.
- The exact shape of the failure region of a structure is typically be unknown at this stage.
- It may still be possible to argue that the failure region belongs to a certain family denoted by $\mathcal{E}$.
- A contour $\partial\mathcal{B}$ will be evaluated with respect to the family $\mathcal{E}$.
- The family $\mathcal{E}$ depends on $\mathcal{B}$ in such a way that $\mathcal{F} \cap \mathcal{B} \subseteq \partial\mathcal{B}$ for all $\mathcal{F} \in \mathcal{E}$.
- If the size of $\mathcal{B}$ is increased (i.e., the structure is strengthened), the family of possible failure regions, $\mathcal{E}$, is reduced, and hence also the failure probability.

Environment contour and failure region



Convex failure regions

It is often natural to assume that a failure region is convex:



This means that if the structure fails at two distinct points (t_1, h_1) and (t_2, h_2), then it also fails for all states on the straight line between these points.

Exceedance probability

The *exceedance probability* of $\mathcal{B}$ with respect to $\mathcal{E}$ is defined as:

$$P_e(\mathcal{B}, \mathcal{E}) = \sup\{P[(T, H) \in \mathcal{F}] : \mathcal{F} \in \mathcal{E}\}.$$

NOTE: The exceedance probability is an upper bound on the failure probability of the structure assuming that the true failure region is a member of the family $\mathcal{E}$.

For a given target exceedance probability $p_e \in (0, 0.5)$ our goal is to find a set $\mathcal{B}$ such that $P_e(\mathcal{B}, \mathcal{E}) = p_e$.

Maximal failure regions

A failure region $\mathcal{F} \in \mathcal{E}$ is said to be *maximal* if there does not exist a region $\mathcal{F}' \in \mathcal{E}$ such that $\mathcal{F} \subset \mathcal{F}'$.

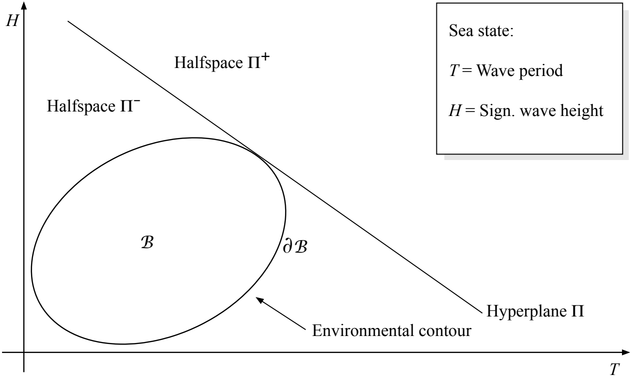
The family of maximal regions in $\mathcal{E}$ is denoted by $\mathcal{E}^*$. If $\mathcal{F}_1, \mathcal{F}_2 \in \mathcal{E}$ and $\mathcal{F}_1 \subseteq \mathcal{F}_2$, we obviously have:

$$P[(T, H) \in \mathcal{F}_1] \leq P[(T, H) \in \mathcal{F}_2].$$

From this it follows that:

$$\begin{aligned} P_e(\mathcal{B}, \mathcal{E}) &= \sup\{P[(T, H) \in \mathcal{F}] : \mathcal{F} \in \mathcal{E}\} \\ &= \sup\{P[(T, H) \in \mathcal{F}] : \mathcal{F} \in \mathcal{E}^*\}. \end{aligned}$$

Supporting halfspaces



- Π is a supporting hyperplane of the convex set $\mathcal{B}$
- Π^+ is a supporting halfspace of the convex set $\mathcal{B}$

Halfspaces and maximal failure regions

Proposition (Halfspace failure region)

Assume that $\mathcal{B}$ is convex and that $\mathcal{E}$ is a family of convex sets such that $\mathcal{F} \cap \mathcal{B} \subseteq \partial\mathcal{B}$ for all $\mathcal{F} \in \mathcal{E}$. We then have:

$$\mathcal{E}^* = \mathcal{P}(\mathcal{B}) = \textit{The family of supporting half-spaces of } \mathcal{B}.$$

Moreover, we have:

$$P_e(\mathcal{B}, \mathcal{E}) = \sup\{P[(T, H) \in \Pi^+] : \Pi^+ \in \mathcal{P}(\mathcal{B})\}.$$

Supporting hyperplane theorems

Theorem (Projection hyperplane)

Let $\mathcal{S} \in \mathbb{R}^n$ be a closed convex set, and assume that $\boldsymbol{x}_0 \notin \mathcal{S}$. Then there exists a supporting hyperplane $\Pi = \{\boldsymbol{x} : \boldsymbol{c}'\boldsymbol{x} = d\}$ of $\mathcal{S}$ such that:

$$\boldsymbol{c}'\boldsymbol{x} \leq d \text{ for all } \boldsymbol{x} \in \mathcal{S}, \quad \text{and} \quad \boldsymbol{c}'\boldsymbol{x}_0 > d.$$

Convex contours

Given a target exceedance probability, $p_e \in (0, 0.5)$ our first aim is to find a convex set $\mathcal{B} \subseteq \mathbb{R}^2$ such that $P_e(\mathcal{B}, \mathcal{E}) = p_e$, where $\mathcal{E}$ is the family of all convex failure regions $\mathcal{F} \subseteq \mathbb{R}^2$ such that $\mathcal{F} \cap \mathcal{B} \subseteq \partial\mathcal{B}$.

We start out by introducing the p_e-level *percentile function* of the joint distribution of (T, H):

$$C(\theta) = \inf\{C : P[T\cos(\theta) + H\sin(\theta) > C] = p_e\}, \quad \theta \in [0, 2\pi).$$

We observe that:

$$C(\theta) = \text{The } (1 - p_e)\text{-percentile of } Y(\theta) = T\cos(\theta) + H\sin(\theta)$$

Convex contours (cont.)

For $\theta \in [0, 2\pi)$ we also introduce :

$$\Pi(\theta) = \{(t, h) : t\cos(\theta) + h\sin(\theta) = C(\theta)\}$$
$$\Pi^+(\theta) = \{(t, h) : t\cos(\theta) + h\sin(\theta) \geq C(\theta)\},$$
$$\Pi^-(\theta) = \{(t, h) : t\cos(\theta) + h\sin(\theta) \leq C(\theta)\}.$$

By the definition of $C(\theta)$ and the assumption that the distribution of (T, H) is absolutely continuous with respect to the Lebesgues measure in $\mathbb{R}^2$ it follows that for all $\theta \in [0, 2\pi)$ we have:

$$P[(T, H) \in \Pi^+(\theta)] = P[T\cos(\theta) + H\sin(\theta) \geq C(\theta)]$$
$$= P[T\cos(\theta) + H\sin(\theta) > C(\theta)] = p_e$$

Proper convex contours

We now assume that there exists a closed convex set $\mathcal{B}$ such that:

$$\mathcal{P}(\mathcal{B}) = \{\Pi^+(\theta) : \theta \in [0, 2\pi)\}. \tag{1}$$

That is, $\Pi^+ \in \mathcal{P}(\mathcal{B})$ if and only if $\Pi^+ = \Pi^+(\theta)$ for some $\theta \in [0, 2\pi)$.

Thus, we may choose any $\Pi^+ \in \mathcal{P}(\mathcal{B})$, and let $\theta \in [0, 2\pi)$ be such that $\Pi^+ = \Pi^+(\theta)$. It then follows that:

$$\begin{aligned} P[(T, H) \in \Pi^+] &= P[(T, H) \in \Pi^+(\theta)] \\ &= P[T\cos(\theta) + H\sin(\theta) > C(\theta)] = p_e. \end{aligned}$$

Since this holds for any $\Pi^+ \in \mathcal{P}(\mathcal{B})$, we must have:

$$P_e(\mathcal{B}, \mathcal{E}) = p_e.$$

If there exists a closed convex set $\mathcal{B}$ such that (1) holds, $\partial\mathcal{B}$ is said to be a *proper convex contour*.

The intersection contour formula

Theorem (Intersection contour)

Assume that there exists a closed convex set $\mathcal{B}$ satisfying (1)*.*

Then $\mathcal{B}$ is given by:

$$\mathcal{B} = \bigcap_{\theta \in [0,2\pi)} \Pi^{-}(\theta).$$

PROOF: By (1) it follows that:

$$\mathcal{B} \subseteq \Pi^{-}(\theta), \text{ for all } \theta \in [0, 2\pi).$$

This implies that:

$$\mathcal{B} \subseteq \bigcap_{\theta \in [0,2\pi)} \Pi^{-}(\theta).$$

The intersection contour formula (cont.)

Assume then that there exists a point $(t_0, h_0) \in \bigcap_{\theta \in [0, 2\pi)} \Pi^-(\theta)$ such that $(t_0, h_0) \notin \mathcal{B}$.

By the projection hyperplane theorem it follows that there exists a hyperplane $\Pi = \{(t, h) : c_1 t + c_2 h = d\}$ such that:

$$c_1 t + c_2 h \leq d \text{ for all } (t, h) \in \mathcal{B},$$
$$c_1 t_0 + c_2 h_0 > d.$$

Without loss of generality we may assume that $\boldsymbol{c} = (c_1, c_2)'$ is a unit vector of the form $(\cos(\theta_0), \sin(\theta_0))'$ for some $\theta_0 \in [0, 2\pi)$.

Thus, the above inequalities may be rewritten as:

$$t \cos(\theta_0) + h \sin(\theta_0) \leq d, \text{ for all } (t, h) \in \mathcal{B},$$
$$t_0 \cos(\theta_0) + h_0 \sin(\theta_0) > d.$$

The intersection contour formula (cont.)

However, since $(t_0, h_0) \in \bigcap_{\theta \in [0,2\pi)} \Pi^-(\theta)$, it follows that we also have:

$$t_0 \cos(\theta_0) + h_0 \sin(\theta_0) \leq C(\theta_0).$$

Hence, by combining these relations we get that:

$$t \cos(\theta_0) + h \sin(\theta_0) < C(\theta_0), \text{ for all } (t, h) \in \mathcal{B},$$

which implies that $\Pi(\theta_0)$ cannot be a supporting hyperplane of $\mathcal{B}$. This contradicts the assumption (1). Hence, we conclude that if (1) holds true, we must have:

$$\mathcal{B} = \bigcap_{\theta \in [0,2\pi)} \Pi^-(\theta). \qquad (2)$$

The intersection contour formula (cont.)

In a given practical situation the C-function must be estimated pointwise using Monte Carlo simulations.

Moreover, $\mathcal{B}$ will typically be approximated by a polygon:

$$\hat{\mathcal{B}} = \bigcap_{i=1}^{n} \hat{\Pi}^{-}(\theta_i),$$

where $\theta_1, \ldots, \theta_n \in [0, \pi)$ are suitably chosen angles.

In the following we shall see how this can be done.

Estimating $C(\theta)$ using Monte Carlo simulation

Assume that we have a sample from the joint distribution of (T, H) generated using Monte Carlo simulation:

$$(T_1, H_1), \dots, (T_n, H_n)$$

For a given angle $\theta \in [0, 2\pi)$ we calculate the projections of these points onto the unit vector $(\cos(\theta), \sin(\theta))$, i.e.:

$$Y_i(\theta) = T_i \cos(\theta) + H_i \sin(\theta), \quad i = 1, \dots, n$$

These projections are then sorted in ascending order:

$$Y_{(1)}(\theta) \leq Y_{(2)}(\theta) \leq \cdots \leq Y_{(n)}(\theta).$$

Estimating $C(\theta)$ using crude Monte Carlo (cont.)

Assuming that $k \leq n$ is an integer such that:

$$\frac{k}{n} \approx 1 - p_e.$$

Then $C(\theta)$ can be estimated by:

$$\hat{C}(\theta) = Y_{(k)}(\theta)$$

Challenges using crude Monte Carlo simulation

- In many typical applications p_e can be very small, i.e., less than 0.1%. In such cases a large number of simulations are needed in order to obtain stable estimates.
- Processing the results in order to obtain the contours can be much more time consuming.
- Storing a large number of simulation results in the computer memory can represent a challenge.
- Most of the simulations yield results close to the central area of the joint distribution, and thus very few results provide information about the contour area.

An improved simulation method will be given later.

Polygon approximation of $\mathcal{B}$

At this stage we assume that we have estimated the C-function for a suitable set of angles $\theta_1, \ldots, \theta_n \in [0, \pi)$, and we let $\hat{C}(\theta_1), \ldots, \hat{C}(\theta_n)$ denote the corresponding estimates. Using these estimates, we define:

$$\hat{\Pi}^+(\theta_i) = \{(t,h) : t\cos(\theta) + h\sin(\theta) \geq \hat{C}(\theta_i)\}$$
$$\hat{\Pi}(\theta_i) = \{(t,h) : t\cos(\theta) + h\sin(\theta) = \hat{C}(\theta_i)\}$$
$$\hat{\Pi}^-(\theta_i) = \{(t,h) : t\cos(\theta) + h\sin(\theta) \leq \hat{C}(\theta_i)\}$$

$\mathcal{B}$ can then be approximated by a polygon of the following form:

$$\hat{\mathcal{B}} = \bigcap_{i=1}^{n} \hat{\Pi}^-(\theta_i).$$

By considering the intersections between consecutive hyperplanes $\Pi(\theta_i)$ and $\Pi(\theta_{i+1})$, we find the corners of the polygon, and hence also the polygon itself.

Polygon approximation of $\mathcal{B}$ (cont.)

Estimating the polygon $\hat{\mathcal{B}}$ using only a few simulations and hyperplanes:

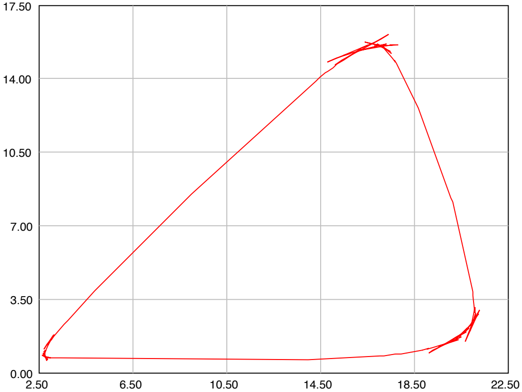


Figure: 1000 simulations, $n = 90$.

Polygon approximation of $\mathcal{B}$ (cont.)

By increasing the number of simulations and hyperplanes, a smoother contour is obtained:

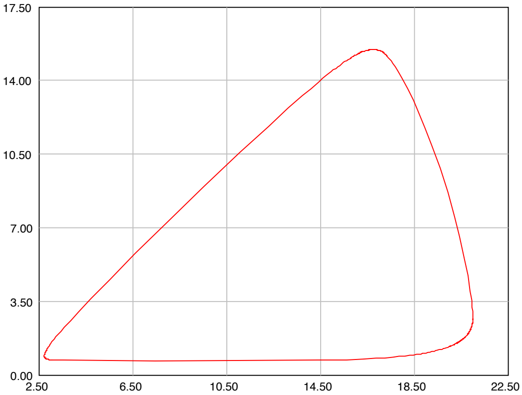


Figure: 1000000 simulations, $n = 360$.

Polygon approximation of $\mathcal{B}$ (cont.)

If we zoom in on the border of $\hat{\mathcal{B}}$, we still find substantial "irregularities":

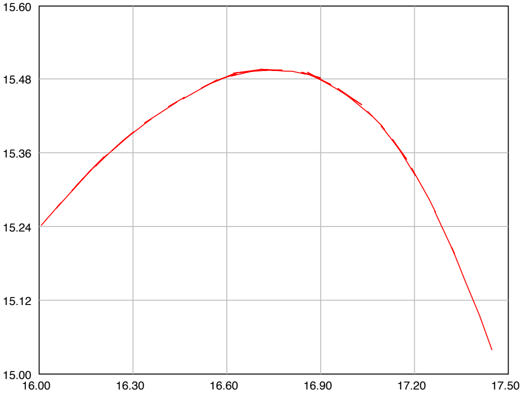


Figure: 1000000 simulations, $n = 360$.

Supporting hyperplanes and contours

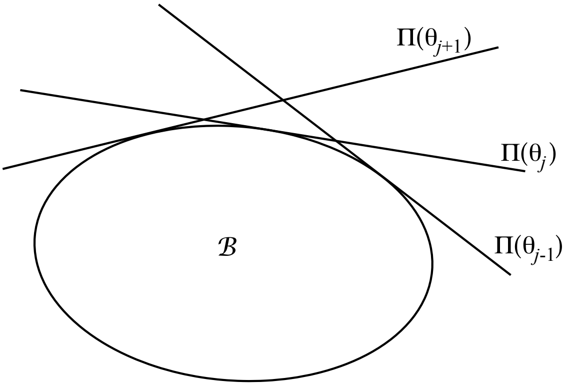


Figure: Ideal case: All hyperplanes support $\mathcal{B}$

Supporting hyperplanes and contours (cont.)

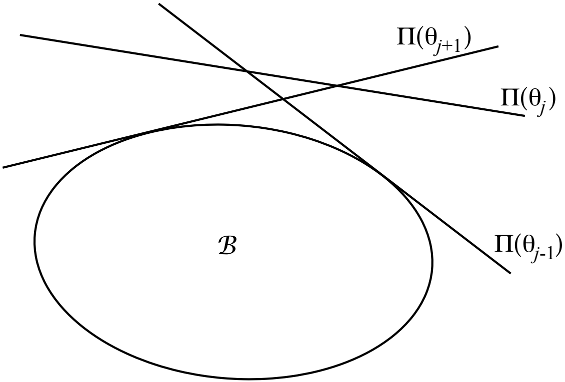


Figure: Irregular case: The hyperplane $\Pi(\theta_j)$ does *not* support $\mathcal{B}$

The "*true*" contour

In the following we assume that $C(\theta)$ is known and differentiable for all $\theta \in [0, 2\pi)$. The boundary $\partial\mathcal{B}$ can then be derived from $C(\theta)$ as follows:

For a given angle $\theta \in [0, 2\pi)$ and a small number $\delta > 0$ consider the intersection between $\Pi(\theta)$ and $\Pi(\theta + \delta)$.

The point (t, h) where the two hyperplanes intersect satisfies:

$$
\begin{aligned}
t\cos(\theta) + h\sin(\theta) &= C(\theta), \\
t\cos(\theta + \delta) + h\sin(\theta + \delta) &= C(\theta + \delta),
\end{aligned}
$$

Multiplying the first equation by $\sin(\theta + \delta)$ and the second equation by $-\sin(\theta)$, we get:

$$
\begin{aligned}
t\cos(\theta)\sin(\theta + \delta) + h\sin(\theta)\sin(\theta + \delta) &= \sin(\theta + \delta)C(\theta), \\
t\cos(\theta + \delta)(-\sin(\theta)) + h\sin(\theta + \delta)(-\sin(\theta)) &= -\sin(\theta)C(\theta + \delta),
\end{aligned}
$$

The "*true*" contour (cont.)

Adding these two equations we get:

$$t[\cos(\theta)\sin(\theta+\delta) - \cos(\theta+\delta)\sin(\theta)] = \sin(\theta+\delta)C(\theta) - \sin(\theta)C(\theta+\delta).$$

We notice that:

$$\cos(\theta)\sin(\theta+\delta) - \cos(\theta+\delta)\sin(\theta) = \sin(\theta+\delta-\theta) = \sin(\delta),$$

and hence, t is given by:

$$t = \frac{\sin(\theta+\delta)C(\theta) - \sin(\theta)C(\theta+\delta)}{\sin(\delta)}.$$

By a similar argument we get that h is given by:

$$h = \frac{-\cos(\theta+\delta)C(\theta) + \cos(\theta)C(\theta+\delta)}{\sin(\delta)}.$$

The "*true*" contour (cont.)

As $\delta \to 0$ the intersection point (t, h) will converge to a point in $\Pi(\theta)$ which we denote by $(t(\theta), h(\theta))$.

Using l'Hôpital's rule we get that:

$$\lim_{\delta \to 0^+} t = t(\theta) = \lim_{\delta \to 0^+} \frac{\cos(\theta + \delta)C(\theta) - \sin(\theta)C'(\theta + \delta)}{\cos(\delta)}$$

$$= C(\theta)\cos(\theta) - C'(\theta)\sin(\theta),$$

and

$$\lim_{\delta \to 0^+} h = h(\theta) = \lim_{\delta \to 0^+} \frac{\sin(\theta + \delta)C(\theta) + \cos(\theta)C'(\theta + \delta)}{\cos(\delta)}$$

$$= C(\theta)\sin(\theta) + C'(\theta)\cos(\theta).$$

The "*true*" contour (cont.)

Hence, $(t(\theta), h(\theta))$ can be written as:

$$\begin{pmatrix} t(\theta) \\ h(\theta) \end{pmatrix} = \begin{bmatrix} C(\theta) & -C'(\theta) \\ C'(\theta) & C(\theta) \end{bmatrix} \cdot \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \tag{3}$$

where $C'(\theta)$ denotes the derivative of $C(\theta)$.

We then let:

$$\partial\mathcal{B} = \{(t(\theta), h(\theta)) : \theta \in [0, 2\pi)\}.$$

Assuming that $\partial\mathcal{B}$ is a *simple closed curve*, $\mathcal{B}$ is the set enclosed by $\partial\mathcal{B}$.

If $\mathcal{B}$ is convex, this implies that $\mathcal{B} \subseteq \Pi^-(\theta)$, and that $\Pi^+(\theta) \cap \mathcal{B} \neq \emptyset$ for all $\theta \in [0, 2\pi)$, and hence it has the correct exceedance probability.

Unfortunately, this may not always be true.

Convexity properties of $\mathcal{B}$

We assume that the $C(\theta)$ is two times differentiable, and consider the derivative of $(t(\theta), h(\theta))$ with respect to θ.

By (3) we get that:

$$
\begin{aligned}
t'(\theta) &= C'(\theta)\cos(\theta) - C(\theta)\sin(\theta) - C''(\theta)\sin(\theta) - C'(\theta)\cos(\theta) \\
&= -[C(\theta) + C''(\theta)]\sin(\theta)
\end{aligned}
$$

$$
\begin{aligned}
h'(\theta) &= C''(\theta)\cos(\theta) - C'(\theta)\sin(\theta) + C'(\theta)\sin(\theta) + C(\theta)\cos(\theta) \\
&= [C(\theta) + C''(\theta)]\cos(\theta).
\end{aligned}
$$

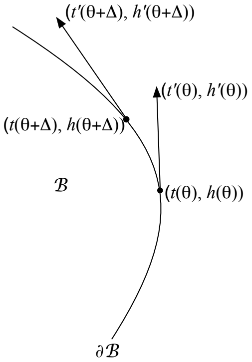
That is, we have:

$$
\begin{pmatrix} t'(\theta) \\ h'(\theta) \end{pmatrix} = [C(\theta) + C''(\theta)] \cdot \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}, \tag{4}
$$

where $C''(\theta)$ denotes the second derivative of $C(\theta)$.

Convexity properties of $\mathcal{B}$ (cont.)

As θ runs through $[0, 2\pi)$, the point $(t(\theta), h(\theta))$ runs counterclockwise through the boundary $\partial\mathcal{B}$.



The derivative $(t'(\theta), h'(\theta))$ is the tangent vector to $\partial\mathcal{B}$ at $(t(\theta), h(\theta))$.

Convexity properties of $\mathcal{B}$ (cont.)

NOTE: Intuitively, the set $\mathcal{B}$ is convex if the angle between $(t'(\theta), h'(\theta))$ and $(t'(\theta + \Delta), h'(\theta + \Delta))$ is positive for any $\theta \in [0, 2\pi)$ and small $\Delta > 0$.

In order to check this, we define:

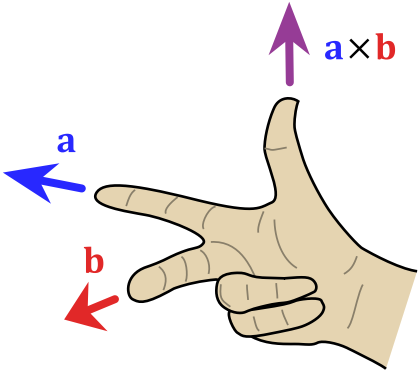
$$\boldsymbol{v}(\theta) = (t'(\theta), h'(\theta), 0), \quad \theta \in [0, 2\pi),$$

and calculate:

$$\boldsymbol{v}(\theta) \times \boldsymbol{v}(\theta + \Delta) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ t'(\theta) & h'(\theta) & 0 \\ t'(\theta + \Delta) & h'(\theta + \Delta) & 0 \end{vmatrix}$$

$$= (0,\ 0,\ t'(\theta) \cdot h'(\theta + \Delta) - h'(\theta) \cdot t'(\theta + \Delta))$$

Convexity properties of $\mathcal{B}$ (cont.)



By the *right-hand rule* of the cross-product the angle between $(t'(\theta), h'(\theta), 0)$ and $(t'(\theta+\Delta), h'(\theta+\Delta), 0)$ is positive if and only if:

$$t'(\theta) \cdot h'(\theta+\Delta) - h'(\theta) \cdot t'(\theta+\Delta) > 0.$$

Convexity properties of $\mathcal{B}$ (cont.)

Inserting the expressions for the derivatives given in (4) we get:

$$
\begin{aligned}
t'(\theta) \cdot h'(\theta + \Delta) - h'(\theta) \cdot t'(\theta + \Delta) \\
&= [C(\theta) + C''(\theta)] \cdot [C(\theta + \Delta) + C''(\theta + \Delta)] \\
&\quad \cdot (-\sin(\theta)\cos(\theta + \Delta) + \sin(\theta + \Delta)\cos(\theta)) \\
&= [C(\theta) + C''(\theta)] \cdot [C(\theta + \Delta) + C''(\theta + \Delta)] \sin(\Delta).
\end{aligned}
$$

Since $\Delta > 0$ is small, we have $\sin(\Delta) > 0$. Hence, the angle between $(t'(\theta), h'(\theta))$ and $(t'(\theta + \Delta), h'(\theta + \Delta))$ is positive if and only if:

$$[C(\theta) + C''(\theta)] \cdot [C(\theta + \Delta) + C''(\theta + \Delta)] > 0, \text{ for all } \theta \in [0, 2\pi), \Delta > 0 \text{ small},$$

By changing the origin of the coordinate system so that $C > 0$, it can be shown that there exists at least one $\theta_0 \in [0, 2\pi)$ such that $C(\theta) + C''(\theta) > 0$. Hence, $\mathcal{B}$ is convex if and only if:

$$C(\theta) + C''(\theta) > 0 \text{ for all } \theta \in [0, 2\pi).$$

Example: A bivariate normal distribution

Assume that T and H are independent normally distributed with $E(T) = \mu_T$, $E(H) = \mu_H$ and $SD(T) = SD(H) = \sigma$.

Thus, $\tilde{T} = (T - \mu_T)/\sigma$ and $\tilde{H} = (H - \mu_H)/\sigma$ are independent standard normally distributed.

By the rotational symmetry property of the standard bivariate normal distribution it follows that:

$$P[\tilde{T}\cos(\theta) + \tilde{H}\sin(\theta) > q_e] = p_e, \text{ for all } \theta \in [0, 2\pi),$$

where q_e denotes the $(1 - p_e)$-percentile of the standard normal distribution.

Hence, for all $\theta \in [0, 2\pi)$ we get that:

$$P[T\cos(\theta) + H\sin(\theta) > \mu_T\cos(\theta) + \mu_H\sin(\theta) + \sigma q_e] = p_e.$$

Convexity properties of $\mathcal{B}$ (cont.)

Thus, we have:

$$C(\theta) = \mu_T \cos(\theta) + \mu_H \sin(\theta) + \sigma q_e,$$
$$C'(\theta) = -\mu_T \sin(\theta) + \mu_H \cos(\theta)$$

By inserting these expressions into (3) we obtain:

$$\begin{aligned} t(\theta) &= C(\theta)\cos(\theta) - C'(\theta)\sin(\theta) \\ &= [\mu_T \cos(\theta) + \mu_H \sin(\theta) + \sigma q_e]\cos(\theta) \\ &- [-\mu_T \sin(\theta) + \mu_H \cos(\theta)]\sin(\theta) \\ &= \mu_T + \sigma q_e \cos(\theta) \end{aligned}$$

$$\begin{aligned} h(\theta) &= C'(\theta)\cos(\theta) + C(\theta)\sin(\theta) \\ &= [-\mu_T \sin(\theta) + \mu_H \cos(\theta)]\cos(\theta) \\ &+ [\mu_T \cos(\theta) + \mu_H \sin(\theta) + \sigma q_e]\sin(\theta) \\ &= \mu_H + \sigma q_e \sin(\theta) \end{aligned}$$

Constructing convex contours (cont.)

Putting it all together we conclude that $\partial\mathcal{B}$ can be written as:

$$\begin{pmatrix} t(\theta) \\ h(\theta) \end{pmatrix} = \begin{pmatrix} \mu_T \\ \mu_H \end{pmatrix} + \begin{bmatrix} \sigma q_e & 0 \\ 0 & \sigma q_e \end{bmatrix} \cdot \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

Moreover, the second derivative of C is given by:

$$C''(\theta) = -\mu_T \cos(\theta) - \mu_H \sin(\theta).$$

Hence, we get:

$$C(\theta) + C''(\theta) = \sigma q_e > 0, \text{ for all } \theta \in [0, 2\pi).$$

In this case $\partial\mathcal{B}$ is a circle with radius σq_e centered at (μ_T, μ_H). Thus, $\partial\mathcal{B}$ is a simple closed curve, and $\mathcal{B}$ is indeed a convex set.

Numerical example based on MC simulation

EXAMPLE: An environmental contour estimated by using Monte Carlo simulation.

We let $p_e = 1.37 \cdot 10^{-5}$, which corresponds to a return period of 25 years.

The joint long-term models for *significant wave height*, denoted by H, and *wave period* denoted by T is given by:

$$f_{T,H}(t,h) = f_H(h) f_{T|H}(t|h)$$

where a three-parameter Weibull distribution is used for the significant wave height, H, and a lognormal conditional distribution is used for the wave period, T.

Numerical example (cont.)

The Weibull distribution is parameterized by a location parameter, γ, a scale parameter α, and a shape parameter β:

$$f_H(h) = \frac{\beta}{\alpha}\left(\frac{h-\gamma}{\alpha}\right)^{\beta-1} e^{-[(h-\gamma)/\alpha]^{\beta}}, \quad h \geq \gamma.$$

The lognormal distribution has two parameters, the log-mean μ and the log-standard deviation σ and is expressed as:

$$f_{T|H}(t|h) = \frac{1}{t\sqrt{2\pi}} e^{-[(\ln(t)-\mu)^2/(2\sigma^2)]}, \quad t \geq 0,$$

Numerical example (cont.)

The dependence between H and T is modelled by letting the parameters μ and σ be expressed in terms of H as follows:

$$\mu = E[\ln(T)|H = h] = a_1 + a_2 h^{a_3},$$

$$\sigma = SD[\ln(T)|H = h] = b_1 + b_2 e^{b_3 h}.$$

The parameters $a_1, a_2, a_3, b_1, b_2, b_3$ are estimated using available data from the relevant geographical location.

Numerical example based on MC simulation (cont.)

TOTAL SEA – WEST OF SHETLAND

Table: Fitted parameter for the three-parameter Weibull distribution for significant wave heights

α	β	γ
2.259	1.285	0.701

Table: Fitted parameter for the conditional log-normal distribution for wave periods

	$i = 1$	$i = 2$	$i = 3$
a_i	1.069	0.898	0.243
b_i	0.025	0.263	-0.148

Numerical example based on MC simulation (cont.)

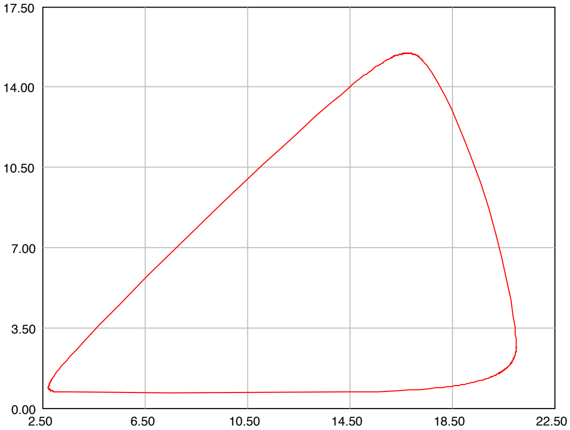
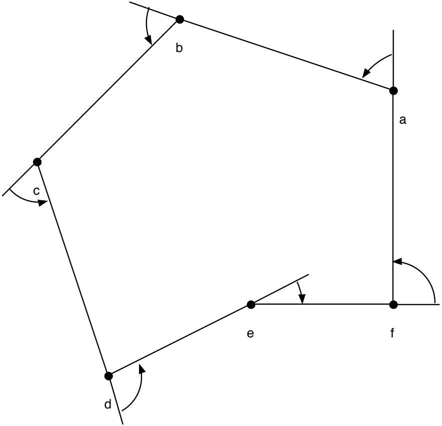


Figure: 1000000 simulations, $n = 360$.

Measuring angles along an environmental contour



Estimated angles along $\partial\mathcal{B}$.

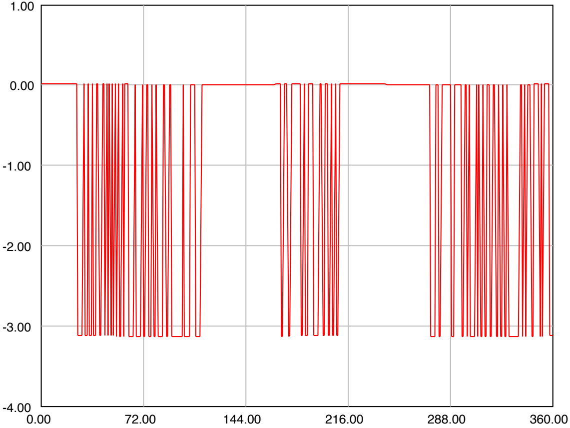


Figure: 1000000 simulations, $n = 360$.

Estimated values for $C(\theta) + C''(\theta)$

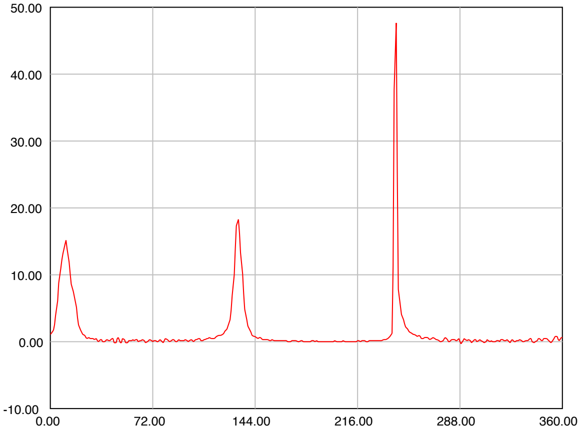


Figure: 1000000 simulations, $n = 360$.

Smoothing the estimated C-curve

To get rid of the loops along the contour, we may use a smoothed version of the estimated C-curve.

$$\tilde{C}(\theta_j) = \frac{\sum_{i=-k}^{+k} \omega_i C(\theta_{j+i})}{\sum_{i=-k}^{+k} \omega_i}, \quad j = 1, \ldots, n,$$

for suitable integer $k \geq 0$ and weights $\omega_{-k}, \ldots, \omega_{+k}$.

NOTE: In the above formula the indices are *"looped"*, so that $\theta_{n+i} = \theta_i$, $i = 1, 2, \ldots, k$, while $\theta_{1-i} = \theta_{n+1-i}$, $i = 1, 2, \ldots, k$.

In the following example we have used $k = 5$ and:

$$\omega_{-i} = \omega_{+i} = (6 - k), \quad i = 0, 1, \ldots, 5.$$

Original (red) and smoothed (green) C-curves

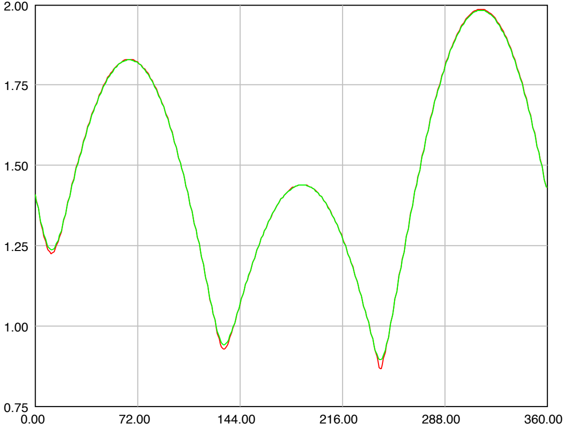


Figure: 1000000 simulations, $n = 360$.

Estimated angles along $\partial\mathcal{B}$ smoothed

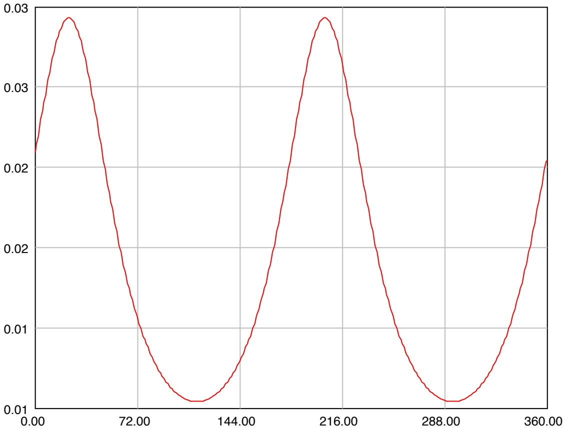


Figure: 1000000 simulations, $n = 360$.

Estimated values for $C(\theta) + C''(\theta)$ smoothed

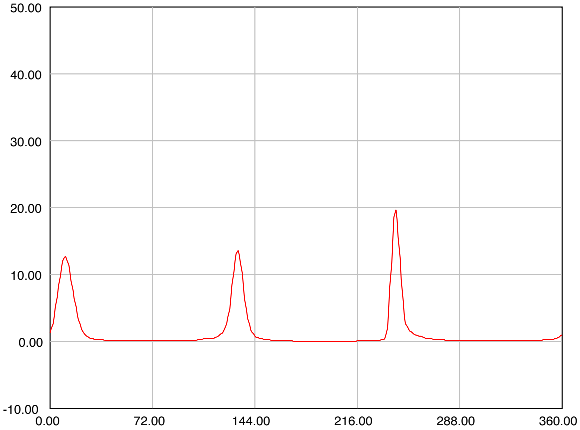


Figure: 1000000 simulations, $n = 360$.

Original and smoothed contours

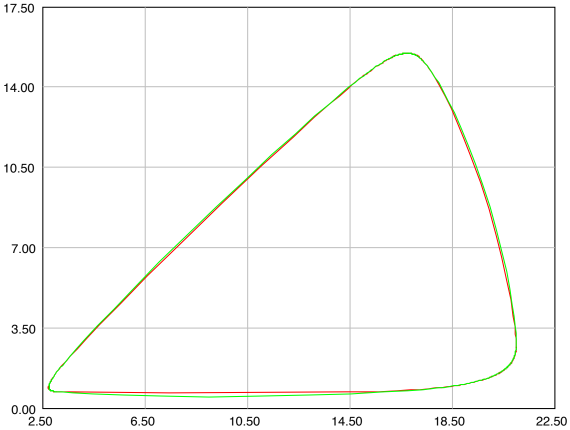


Figure: 1000000 simulations, $n = 360$.

Transformed contours

Transformed contours

Transformed contours

The *Rosenblatt transformation*, denoted Ψ, depends on the joint distribution of (T, H), and is such that if:

$$(T', H') = \Psi(T, H),$$

then T' and H' are independent standard normally distributed.

- The *normal space* = The space containing (T', H').
- The *environmental space* = The space containing (T, H).

For a given set $\mathcal{B}'$ in the *normal space*, we let $\mathcal{E}'$ be the family of all convex sets $\mathcal{F}'$ in the *normal space* such that $\mathcal{F}' \cap \mathcal{B}' \subseteq \partial\mathcal{B}'$.

If $\mathcal{B}'$ is convex, it follows by the halfspace failure region proposition that:

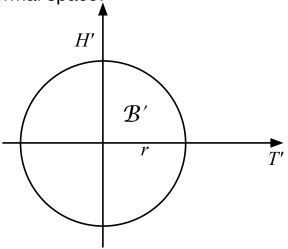
$$\mathcal{E}'^* = \mathcal{P}(\mathcal{B}').$$

Transformed contours (cont.)

Let $p_e < 0.5$ be the desired exceedance probability, and let $r > 0$ denote the $(1 - p_e)$-percentile in the standard normal distribution.

A contour $\partial\mathcal{B}'$ for (T', H') is constructed by letting $\mathcal{B}'$ be a circle centered at the origin and with radius r.

Normal space:



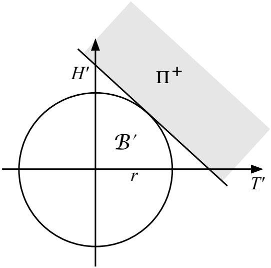
NOTE: Since T' and H' are standard normally distributed, it follows that:

$$P[T' > r] = P[H' > r] = p_e.$$

Transformed contours (cont.)

If $\Pi^+ \in \mathcal{P}(\mathcal{B}')$, it follows by the rotational symmetry property of normal distribution that:

$$P[(T', H') \in \Pi^+] = P[T' > r] = p_e.$$



Since this is true for all $\Pi^+ \in \mathcal{P}(\mathcal{B}')$, we then get:

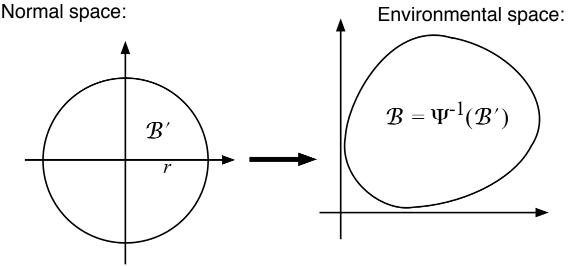
$$P_e(\mathcal{B}', \mathcal{E}') = \sup\{P[(T', H') \in \Pi^+] : \Pi^+ \in \mathcal{P}(\mathcal{B}')\} = p_e.$$

Transformed contours (cont.)

The set $\mathcal{B}$ is then obtained by transforming the set $\mathcal{B}'$ from normal space back to the environmental space using the *inverse Rosenblatt transformation*.

That is, we let $\mathcal{B}$ be given by:

$$\mathcal{B} = \Psi^{-1}(\mathcal{B}') = \{(t, h) = \Psi^{-1}(x, y) : (x, y) \in \mathcal{B}'\}$$



The inverse Rosenblatt transformation

The transformation of (T', H') into (T, H) is done in two steps:

Step 1. Transform (T', H') into (U, V) such that U and V are independent and uniformly distributed on $[0, 1]$.

Step 2. Transform (U, V) into (T, H)

We let Φ denote the cumulative distribution function of the standard normal distribution. Thus, if X is standard normally distributed, we have:

$$P(X \leq x) = \Phi(x).$$

The cumulative distribution function of H is denoted F_H, while the conditional distribution function of T given H is denoted $F_{T|H}$.

The inverse Rosenblatt transformation (cont.)

Step 1. In this step we let:

$$U = \Phi(H'),$$

$$V = \Phi(T').$$

This implies that for all $u, v \in [0, 1]$ we have:

$$\begin{aligned} P(U \leq u) &= P(\Phi(H') \leq u) = P(H' \leq \Phi^{-1}(u)) \\ &= \Phi(\Phi^{-1}(u)) = u, \end{aligned}$$

$$\begin{aligned} P(V \leq v) &= P(\Phi(T') \leq v) = P(T' \leq \Phi^{-1}(v)) \\ &= \Phi(\Phi^{-1}(v)) = v. \end{aligned}$$

Hence, U and V are independent and uniformly distributed on $[0, 1]$.

The inverse Rosenblatt transformation (cont.)

Step 2. In this step we let:

$$H = F_H^{-1}(U),$$

$$T = F_{T|H}^{-1}(V|H).$$

This implies that for all h, t we have:

$$\begin{aligned} P(H \leq h) &= P(F_H^{-1}(U) \leq h) \\ &= P(U \leq F_H(h)) = F_H(h), \end{aligned}$$

$$\begin{aligned} P(T \leq t|H = h) &= P(F_{T|H}^{-1}(V|h) \leq t|H = h) \\ &= P(V \leq F_{T|H}(t|h)|H = h) = F_{T|H}(t|h). \end{aligned}$$

Mixtures of distributions

Handling discrete mixtures of distributions

Mixtures of distributions

Mixtures of distributions occur when modelling environmental variables which depend on different background variables:

- Seasonal effects – Separate models fitted for each season (or month)
- Directional effects – Separate models fitted for each direction (north, east, south and west)

NOTE: Only discrete mixtures will be considered here.

Mixtures of distributions (cont.)

We focus on the distribution of H only. The conditional distribution of $T|H$ is handled completely similar.

Assume that the cumulative distribution function for H is a the mixture of $F_{H,1}, \ldots, F_{H,m}$:

$$F_H(h) = \sum_{j=1}^{m} \alpha_j F_{H,j}(h),$$

where $\alpha_j \geq 0$, $j = 1, \ldots, m$, and $\sum_{j=1}^{m} \alpha_j = 1$.

We assume that $F_{H,1}, \ldots, F_{H,m}$ are m cumulative distribution functions which are all continuous and strictly increasing. Moreover, we assume that the inverse functions $F_{H,1}^{-1}, \ldots, F_{H,m}^{-1}$ are known and easy to calculate.

Mixtures of distributions (cont.)

Assume that $U = u$, and that we want to compute H, $h = F_H^{-1}(u)$. This is equivalent to solving the following equation:

$$F_H(h) = \sum_{j=1}^{m} \alpha_j F_{H,j}(h) = u.$$

We claim that if h is the solution to this equation, then:

$$h_{min} = \min_{1 \leq j \leq m} h_j \leq h \leq h_{max} = \max_{1 \leq j \leq m} h_j,$$

where:

$$h_j = F_{H,j}^{-1}(u), \quad j = 1, \ldots, m.$$

Mixtures of distributions (cont.)

To prove this we note that since the cumulative distribution functions are non-decreasing and $\sum_{j=1}^{m} \alpha_j = 1$, we have:

$$F_H(h_{min}) = \sum_{j=1}^{m} \alpha_j F_{H,j}(h_{min}) \leq \sum_{j=1}^{m} \alpha_j F_{H,j}(h_j) = \sum_{j=1}^{m} \alpha_j u = u$$

Similarly, we have:

$$F_H(h_{max}) = \sum_{j=1}^{m} \alpha_j F_{H,j}(h_{max}) \geq \sum_{j=1}^{m} \alpha_j F_{H,j}(h_j) = \sum_{j=1}^{m} \alpha_j u = u$$

Since $F_{H,1}, \ldots, F_{H,m}$ are continuous and strictly increasing, it follows that F_H is continuous and strictly increasing as well. Thus, since:

$$F_H(h_{min}) \leq u \leq F_H(h_{max})$$

there must exist some $h \in [h_{min}, h_{max}]$ such that $F_H(h) = u$.

Mixtures of distributions (cont.)

Having identified the interval $[h_{min}, h_{max}]$ which must contain a unique h such that:

$$F_H(h) = \sum_{j=1}^{m} \alpha_j F_{H,j}(h) = u.$$

the solution to this equation can easily be found numerically, e.g., by using the *bisection method*.

Exceedance probability of transformed contours

Estimating the exceedance probability of transformed contours

Convex failure regions in the normal space

In the environmental space we may compute the exceedance probability of the contour for the case where $\mathcal{E}$ is given by:

$$\mathcal{E} = \{\mathcal{F} = \Psi^{-1}(\mathcal{F}') : \mathcal{F}' \in \mathcal{E}'\},$$

If we do this, we get:

$$\begin{aligned} P_e(\mathcal{B}, \mathcal{E}) &= \sup\{P[(T,H) \in \mathcal{F}] : \mathcal{F} \in \mathcal{E}\} \\ &= \sup\{P[(T,H) \in \Psi^{-1}(\mathcal{F}')] : \mathcal{F}' \in \mathcal{E}'\} \\ &= \sup\{P[(T',H') \in \mathcal{F}'] : \mathcal{F}' \in \mathcal{E}'\} \\ &= \sup\{P[(T',H') \in \mathcal{F}'] : \mathcal{F}' \in \mathcal{E}'^*\} \\ &= \sup\{P[(T',H') \in \Pi^+] : \Pi^+ \in \mathcal{P}(\mathcal{B}')\} = p_e. \end{aligned}$$

Hence, the contour $\partial\mathcal{B}$ has the desired exceedance probability with respect to the family $\mathcal{E}$ of failure regions.

Exceedance probability of transformed contours

PROBLEM: We observe that $\mathcal{E}$ consists of transformed convex sets, where the transformation depends on the joint distribution of (T, H).

Environmental conditions may vary a lot from location to location. If the family $\mathcal{E}$ depends on the joint distribution of (T, H), $\mathcal{E}$ also varies from location to location.

The true failure region of a given mechanical construction, however, should be the same irrespective of location.

In the following we instead assume that $\mathcal{E}$ is the family of convex sets $\mathcal{F}$ such that $\mathcal{F} \cap \mathcal{B} \subseteq \partial\mathcal{B}$.

Convex failure regions in the environmental space

We recall that the *exceedance probability* of $\mathcal{B}$ with respect to $\mathcal{E}$ is defined as:

$$P_e(\mathcal{B}, \mathcal{E}) = \sup\{P[(T, H) \in \mathcal{F}] : \mathcal{F} \in \mathcal{E}\}.$$

Since a transformed set $\mathcal{B}$ may not itself be convex, the family of maximal failure regions, $\mathcal{E}^*$ is in general not equal to $\mathcal{P}(\mathcal{B})$. Hence, it turns out to be difficult to go through all sets $\mathcal{F} \in \mathcal{E}^*$ in order to identify the set with the highest probability.

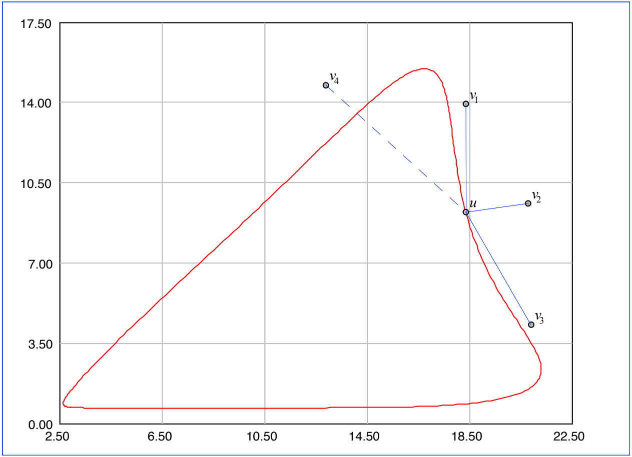
Instead we work with a slightly modified family of failure regions denoted $\tilde{\mathcal{E}}$, defined as follows:

$$\tilde{\mathcal{E}} = \{\tilde{\mathcal{F}}(u) : u \in \partial\mathcal{B}\},$$

where $\tilde{\mathcal{F}}(u)$ is the set of all points $v \notin \mathcal{B}$ that are *visible* from u.

A point $v \notin \mathcal{B}$ is said to be visible from u if the line between u and v does not intersect the interior of $\mathcal{B}$.

Visible points



From the point u, the three points v_1, v_2 and v_3 are visible, while the point v_4 is *not* visible.

Visibility method

Proposition

Assume that $\mathcal{F} \in \mathcal{E}$. Then there exists a set $\tilde{\mathcal{F}} \in \tilde{\mathcal{E}}$ such that $\mathcal{F} \subseteq \tilde{\mathcal{F}}$.

PROOF: Let $\mathcal{F} \in \mathcal{E}$. Then there exists a set $\mathcal{F}^* \in \mathcal{E}^*$ such that $\mathcal{F} \subseteq \mathcal{F}^*$. Then there must exist at least one point $u \in \partial\mathcal{B}$ such that $u \in \mathcal{F}^*$. If this is not the case, this contradicts that $\mathcal{F}^*$ is a maximal failure region.

Now, let $v \in \mathcal{F}^*$ be arbitrary. Since $\mathcal{F}^*$ is convex, the line segment between u and v is contained inside $\mathcal{F}^*$.

Since $\mathcal{F}^* \cap \mathcal{B} \subseteq \partial\mathcal{B}$, the line segment between u and v does not intersect the interior of $\mathcal{B}$.

Hence, v is visible from u, and since v was chosen arbitrarily, this implies that all points in $\mathcal{F}^*$ are visible from u.

Thus, by letting $\tilde{\mathcal{F}} = \tilde{\mathcal{F}}(u) \in \tilde{\mathcal{E}}$, we get that:

$$\mathcal{F} \subseteq \mathcal{F}^* \subseteq \tilde{\mathcal{F}}.$$

Visibility method (cont.)

Corollary

Let $\mathcal{B}$ be an environmental contour set. Then we have:

$$P_e(\mathcal{B}, \mathcal{E}) \leq P_e(\mathcal{B}, \tilde{\mathcal{E}})$$

PROOF: Let $\mathcal{F} \in \mathcal{E}$. Then by the above proposition there exists a set $\tilde{\mathcal{F}} \in \tilde{\mathcal{E}}$ such that $\mathcal{F} \subseteq \tilde{\mathcal{F}}$.

Hence, we have:

$$P[(T, H) \in \mathcal{F}] \leq P[(T, H) \in \tilde{\mathcal{F}}]$$

From this it follows that:

$$\begin{aligned} P_e(\mathcal{B}, \mathcal{E}) &= \sup\{P[(T, H) \in \mathcal{F}] : \mathcal{F} \in \mathcal{E}\} \\ &\leq \sup\{P[(T, H) \in \tilde{\mathcal{F}}] : \tilde{\mathcal{F}} \in \tilde{\mathcal{E}}\} \\ &= P_e(\mathcal{B}, \tilde{\mathcal{E}}). \end{aligned}$$

Visible points - example 1

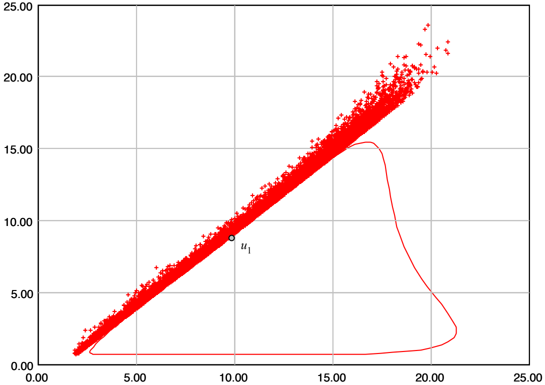


Figure: Estimation of $P((T, H) \in \tilde{\mathcal{F}}(u_1))$ using simulated visible points

Visible points - example 2

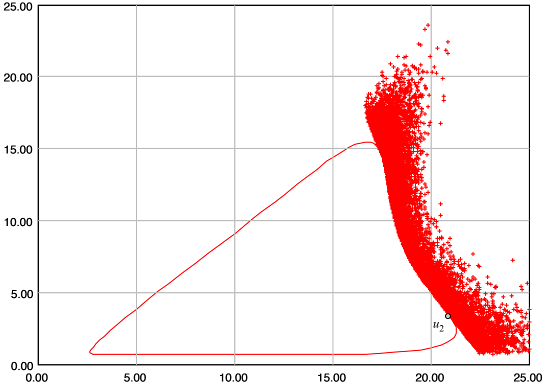


Figure: Estimation of $P((T, H) \in \tilde{\mathcal{F}}(u_2))$ using simulated visible points

Visibility method (cont.)

It can be shown that if $\mathcal{B}$ is *convex* and $\partial\mathcal{B}$ is a differentiable curve, we get that:

$$P_e(\mathcal{B}, \mathcal{E}) = P_e(\mathcal{B}, \tilde{\mathcal{E}}).$$

Thus, in such cases $P_e(\mathcal{B}, \tilde{\mathcal{E}})$ is exact. Furthermore, in general the sets in $\tilde{\mathcal{E}}$ are *almost convex*. Thus, $P_e(\mathcal{B}, \tilde{\mathcal{E}})$ is typically a very good upper bound on the true exceedance probability.

An efficient algorithm for estimating $P_e(\mathcal{B}, \tilde{\mathcal{E}})$ is given in Huseby et al (2019).

Numerical example

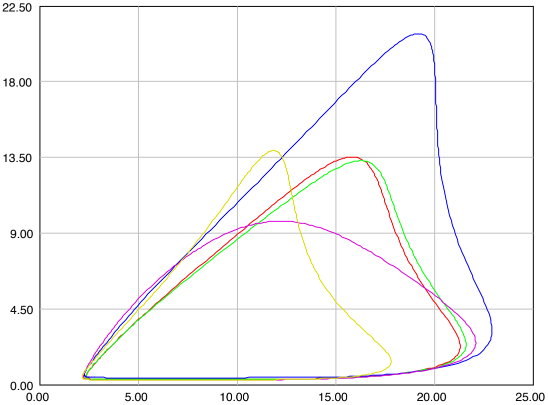


Figure: Environmental contours for five individual submodels before mixing.

Numerical example (cont.)

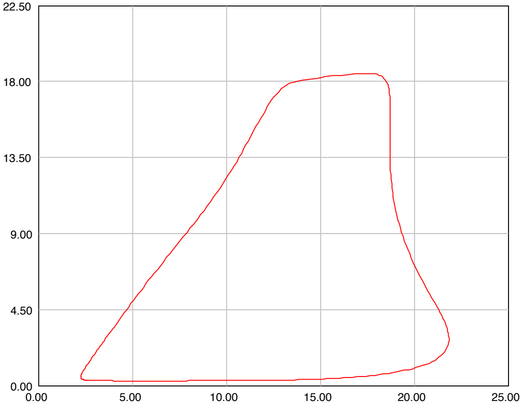


Figure: Environmental contour for the mixed bivariate distribution.

Numerical example (cont.)

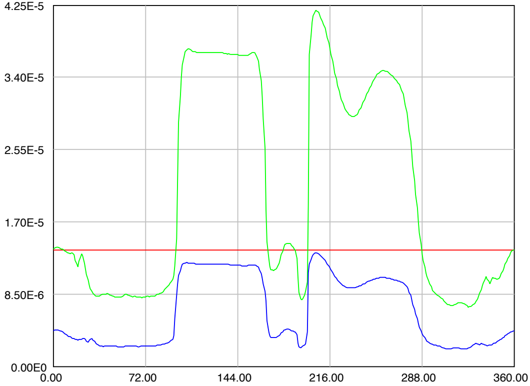


Figure: Estimated probabilities for the sets $\tilde{\mathcal{F}}(u_i)$, $i = 1, \ldots, 360$ for the original environmental contour (green), the desired exceedance probability (red) and an adjusted environmental contour (blue)

Numerical example (cont.)

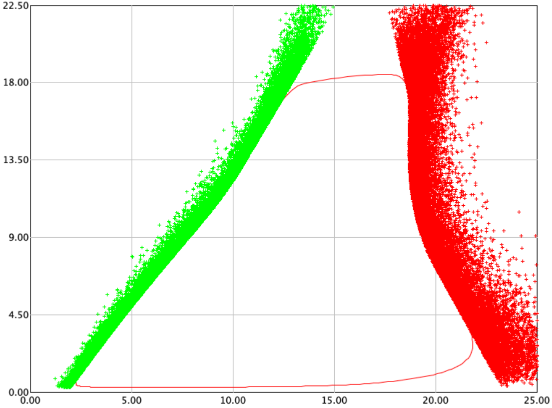


Figure: Environmental contour along with simulated outcomes in the sets $\tilde{\mathcal{F}}(u_{105})$ (red scatter) and $\tilde{\mathcal{F}}(u_{205})$ (green scatter).

Numerical example (cont.)

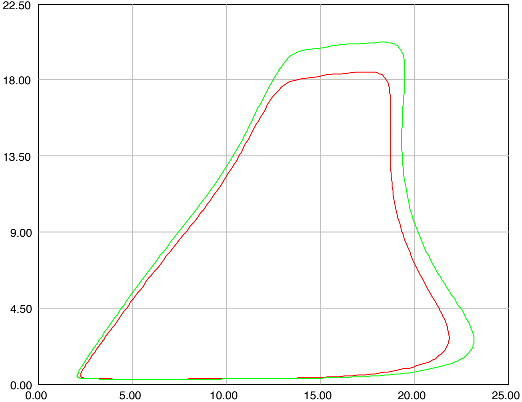


Figure: Original (red curve) and adjusted (green curve) environmental contours for the mixed bivariate distribution.

Estimating $C(\theta)$ revisited

Improved estimation method

Estimating $C(\theta)$ revisited

Assume that X and Y are independent and normally distributed with mean 0 and standard deviation 1. We then let:

$$R = \sqrt{X^2 + Y^2}, \\ V = \mathrm{atan2}(Y, X),$$

where the function $\mathrm{atan2}(y, x)$ is defined as follows:

$$\mathrm{atan2}(y, x) = \begin{cases} \arctan(\frac{y}{x}) & \text{if } x > 0, \\ \arctan(\frac{y}{x}) + \pi & \text{if } x < 0 \text{ and } y \geq 0, \\ \arctan(\frac{y}{x}) - \pi & \text{if } x < 0 \text{ and } y < 0, \\ \frac{\pi}{2} & \text{if } x = 0 \text{ and } y > 0, \\ -\frac{\pi}{2} & \text{if } x = 0 \text{ and } y < 0, \\ \text{undefined} & \text{if } x = 0 \text{ and } y = 0. \end{cases}$$

This implies that R and V are the polar coordinates of (X, Y).

Estimating $C(\theta)$ revisited (cont.)

Now R and V are independent, and $Z = R^2$ is χ^2-distributed with 2 degrees of freedom, while V is $R[0, 2\pi]$-distributed.

This means that the density of Z is :

$$f_Z(z) = \frac{1}{2} e^{-z/2}, \quad \text{for } z > 0,$$

which is an exponential distribution with rate $\lambda = 1/2$.

This implies that $P(Z > z) = e^{-z/2}$.

Thus, the probability that (X, Y) is located outside a circle with centrum in origin and with a radius r is equal to $e^{-r^2/2}$.

Estimating $C(\theta)$ revisited (cont.)

To simulate from the distribution of (X, Y) we start by generating U and V, where $U \sim R[0, 1]$ and $V \sim R[0, 2\pi]$.

We then let $Z = -2\ln(U)$. Now, it is easy to show that Z gets the density f_Z. We also calculate $R = \sqrt{Z}$. Since R and V are the polar coordinates to (X, Y), we find that:

$$X = R\cos(V) = \sqrt{Z} \cdot \cos(V),$$
$$Y = R\sin(V) = \sqrt{Z} \cdot \sin(V).$$

We then let $(T, H) = \Psi^{-1}(X, Y)$, where Ψ^{-1} is the inverse Rosenblatt transformation for the joint distributions of T and H. This way (T, H) gets the correct joint distribution.

Estimating $C(\theta)$ revisited (cont.)

Let $\theta \in [0, 2\pi)$, and let $S(\theta) = T\cos(\theta) + H\sin(\theta)$.

For a given exceedance probability p_e we wish to estimate $C(\theta)$ such that $P(S(\theta) > C(\theta)) = p_e$.

By simulating (T, H) n times, each time calculating the resulting value of $S(\theta)$, we can estimate $C(\theta)$ by the order observator $S_{(k)}(\theta)$, where k is such that:

$$1 - \frac{k}{n} = \frac{n-k}{n} \approx p_e.$$

If p_e is very small, i.e., 0.1%, a large number of simulations are needed in order to obtain stable estimates.

Most of the simulations yield results close to the central area of the joint distribution. Very few of the simulated values provide information about the contour area.

Alternative Monte Carlo sampling scheme

IDEA: Avoid sampling points from the central area of the joint distribution, and just sample points close to the contour.

Simulate (X, Y) from the conditional distribution for (X, Y) given that this vector falls outside a circle with radius, say r_0.

Simulate (X, Y) from the conditional distribution given that $R = \sqrt{X^2 + Y^2} > r_0$.

Simulate (X, Y) from the conditional distribution for (X, Y) given that $Z = X^2 + Y^2 > z_0 = r_0^2$.

Alternative Monte Carlo sampling scheme (cont.)

The conditional distribution for Z given that $Z > z_0$ is given by:

$$P(Z > z | Z > z_0) = \frac{P(Z > z \bigcap Z > z_0)}{P(Z > z_0)} = \frac{P(Z > z)}{P(Z > z_0)} = e^{-(z-z_0)/2}.$$

Hence, given that $Z > z_0$, $(Z - z_0)$ is exponentially distributed with $\lambda = 1/2$.

Thus, we can simulate from the conditional distribution for Z given $Z > z_0$ by generating $U \sim R[0, 1]$ and let:

$$Z = z_0 - 2\ln(U) = r_0^2 - 2\ln(U).$$

The angle V is generated from the $R[0, 2\pi]$-distribution.

Finally, we let $X = \sqrt{Z}\cos(V)$ and $Y = \sqrt{Z}\sin(V)$, and $(T, H) = \Psi^{-1}(X, Y)$, where Ψ^{-1} is the inverse Rosenblatt transformation.

Alternative Monte Carlo sampling scheme (cont.)

NOTE: The simulations are focused in the area of interest on the outer edge of the outcome space where we expect that the contour is.

However, we need to correct for this by estimating the percentile function $C(\theta)$ using an *adjusted* exceedance probability which takes into acount that we are not simulating from the true joint distributions of T and H.

We let $p_e' = P(S(\theta) > C(\theta) | R > r_0)$ be this adjusted exceedance probability, and assume that r_0 is chosen such that the event $\{S(\theta) > C(\theta)\}$ is contained in the event $\{R > r_0\}$.

We can achieve this by ensuring that r_0 is *not too large*.

Alternative Monte Carlo sampling scheme (cont.)

Assuming that the event $\{S(\theta) > C(\theta)\}$ is contained in the event $\{R > r_0\}$, we have:

$$p'_e = P(S(\theta) > C(\theta) | R > r_0) = \frac{P(S(\theta) > C(\theta) \bigcap R > r_0)}{P(R > r_0)}$$

$$= \frac{P(S(\theta) > C(\theta))}{P(R > r_0)}$$

$$= \frac{p_e}{e^{-r_0^2/2}} = e^{r_0^2/2} \cdot p_e,$$

where we have used that:

$$P(R > r_0) = P(Z > r_0^2) = e^{-r_0^2/2}.$$

Alternative Monte Carlo sampling scheme (cont.)

We can then simulate n times from this conditional distribution and estimate $C(\theta)$ by the order observation $S_{(k')}(\theta)$, but where k' is determined so that:

$$1 - \frac{k'}{n} = \frac{n - k'}{n} \approx p'_e = e^{r_0^2/2} \cdot p_e.$$

NOTE: Since $r_0 > 0$, we have $e^{r_0^2/2} > 1$.

Hence, $p'_e > p_e$ and $k' < k$. This means that a (much) larger fraction of the simulated data is used to estimate $C(\theta)$.

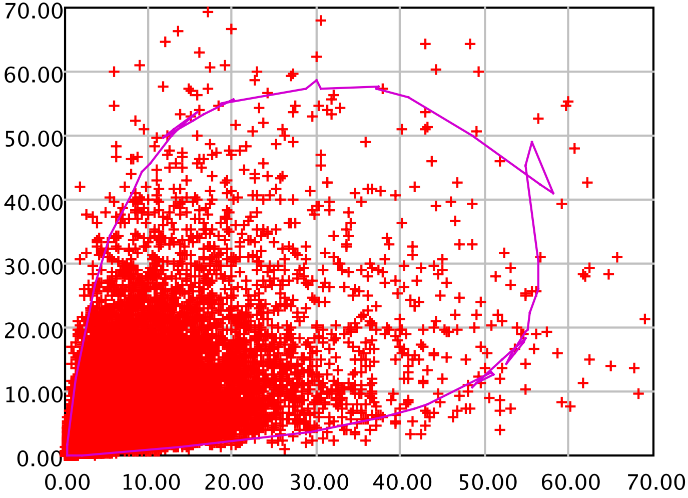
Alternative Monte Carlo sampling scheme (cont.)

Ideally, we would like r_0 to be as large as possible to maximize the effect of the importance sampling. At the same time we must ensure that the event $\{S(\theta) > C(\theta)\}$ is contained in the event $\{R > r_0\}$.

We let $\mathcal{O}$ denote a circle centered in the origin with radius r_0. Then r_0 must be chosen so that the transformed set $\Psi^{-1}(\mathcal{O})$ is contained inside the contour we want to estimate.

Experiences has shown that we get a stable estimate by choosing $r_0 = 0.95 \cdot r$, where r is the radius we use to determine the transformed contour.

Estimating a contour *without* importance sampling



Estimating a contour *with* importance sampling