Decision-theoretic reliability sensitivity

Daniel Straub, Max Ehre, Iason Papaioannou

Engineering Risk Analysis Group,
Technische Universität München, Germany

Abstract

We propose and discuss sensitivity metrics for reliability analysis, which are based on the value of information. These metrics are easier to interpret than other existing sensitivity metrics in the context of a specific decision and they are applicable to any type of reliability assessment, including those with dependent inputs. We develop computational strategies that enable efficient evaluation of these metrics, in some scenarios without additional runs of the deterministic model. The metrics are investigated by application to numerical examples.

Keywords: Sensitivity analysis, value of information, reliability

1. Introduction

The evaluation of sensitivities is an essential part of scientific and engineering analysis. Sensitivities provide information on the relative importance of model input quantities and they support optimization as well as model checking. A large number of sensitivity metrics have been proposed in the literature [1–3]. Metrics exist for deterministic model inputs and for random model inputs. One can distinguish local vs. global metrics. Among the latter, there are variance-based [4–6] or distribution-based [7–9] metrics.

The proper choice of a sensitivity metric depends on the application and decision context. In this contribution, we consider sensitivity analysis for random input quantities $X$ to models that are utilized to evaluate the reliability of a system. Formally, we consider a model $g(X)$, whose output describes the performance of the system in such a way that a value of $g(X)$ below some threshold corresponds to failure $F$. Without loss of generality, this threshold can be set to zero, so that failure is $F = \{g(X) \leq 0\}$ and the probability of failure is $p_F = \Pr[g(X) \leq 0]$. In structural reliability, $g$ is called limit-state function (LSF); in system reliability, $g$ is the structure function.

Many authors have proposed metrics for such a reliability sensitivity analysis [e.g., 10–17]. The sensitivity can be with respect to deterministic input parameters $\theta$, including distribution parameters. In this case, the probability of failure $p_F$ can be interpreted as a function of $\theta$ and a common sensitivity metric is the partial derivative of $p_F$ with respect to $\theta$, $\frac{\partial p_F(\theta)}{\partial \theta_i}$ [12, 13, 18].

In many instances, the interest is in the sensitivity with respect to the input random variables $X$. A popular metric in this context are the FORM (First-Order Reliability Method) $\alpha$-factors [10, 14], or derivatives thereof [11, 13–15]. The FORM $\alpha$-factors are briefly reviewed in Section 3.3. As an alternative, classical global sensitivity metrics, specifically variance-based and distribution-based metrics, have been adapted for reliability analyses [19, 22]. In addition, quantile-based sensitivity metrics can be utilized in the context of reliability and risk analysis [23].

In many application and decision contexts, the motivation for the sensitivity analysis is to understand on which input random variables one should collect more information to reduce their uncertainty. One refers to this setting as Factor Prioritization [1]. To provide a meaningful prioritization, one needs to understand the purpose of the model. In an engineering context, models ultimately serve as decision support. Hence the prioritization should consider how the uncertainty in the input random variables affects the decisions taken based on the model outcome, i.e., by a measure of decision sensitivity [24]. In decision analysis, the
The value of information concept quantifies the effect of reducing uncertainty on the optimality of a decision made based on the model outcome \[25, 26\]. The use of value of information for sensitivity analysis dates back to Felli and Hazen \[24\], who proposed the expected value of perfect information \(1\) (EVPPI) as a sensitivity metric in the context of medical decision making. Around the same time, Pörn \[27\] suggested the use of EVPPI as a component importance measure in Probabilistic Safety Assessments (PSA). More recently, also Borgonovo and Cillo \[28\], Fauriat and Zio \[29\] and Bjørnsen et al. \[30\] propose and demonstrate value of information concepts for sensitivity analysis for PSA. In PSA, the system is a deterministic function of binary components with availabilities \(p_i\), e.g., through a fault tree model; i.e., the \(X\) are Bernoulli random variables and \(g\) is the so-called structure function \[31\].

In contrast to these works, we focus on the application of value of information for reliability sensitivity when the reliability is estimated based on a general model with discrete or continuous input random variables. Examples are assessments using physics-based system performance models, such as those found in structural reliability applications \[13, 32\]. In this context, the value of information concept is applied mainly for optimization of inspection and monitoring schemes \[e.g., 33–39\]. Rather surprisingly, its use as a sensitivity metric has received little attention.

In this paper, we develop the EVPPI sensitivity metric for two key decision contexts. The first one corresponds to a safety assessment case, in which the decision is between accepting the current system or performing an upgrade that significantly increases reliability. The second decision context corresponds to a reliability-based design, in which one or more design parameters can be selected from a typically continuous set of alternatives. The definition of the EVPPI for these two decision contexts, together with algorithms for computing them, are presented in Sections 3 and 4 after a general introduction to the value of information in Section 2. We also demonstrate the relation of the proposed reliability sensitivity metrics to existing sensitivity metrics, in particular the FORM sensitivities and reliability-oriented variance-based sensitivity metrics.

Computation of the EVPPI can be computationally demanding, in particular if the analysis involves advanced computer models. This has been addressed by Oakley and co-authors \[40–42\], who present computation strategies for evaluating the EVPPI as well as the expected value of sample information (EVSI) through the use of surrogate modeling strategies. In this paper, we present computational strategies to efficiently evaluate the EVPPI in reliability applications for the two decision contexts. In Section 3, we show that for the safety assessment case, the EVPPI can be obtained by a mere post-processing of the reliability analysis results without additional runs of the model \(g\).

The proposed sensitivity metrics and their computation are investigated and demonstrated by application to examples in Section 5.

2. Value of information

The ultimate goal of an engineering analysis is the recommendation of an optimal action or decision. Example decisions are the selection of a system design, the choice of a management strategy or the decision on whether or not an engineering system is safe to be operated. In line with the literature on decision analysis \[25, 43\], \(a\) denotes a decision alternative.

Following expected utility theory, the optimal decision \(a_{opt}\) is the one maximizing the expected utility or – equivalently when the utility is linear with loss – minimizing the expected loss \(L\):

\[
a_{opt} = \arg \min_a E_X [L(X, a)] .
\]

\(X\) is the vector of random variables affecting the loss and \(E_X\) is the mathematical expectation with respect to the probability measure of \(X\).

---

\(^1\)Felli and Hazen, as well as most other authors, refer to the metric as expected value of perfect information (EVPI). In agreement with the general decision analysis literature we use the term EVPPI for the same metric. The reason for this clarification becomes apparent in Section 2.
The loss function $L$ is problem specific. A frequently used loss function is of the quadratic form,

$$L(\mathbf{X}, a) = [Y(\mathbf{X}) - a]^2.$$  

(2)

This loss function occurs when the system design $a$ is optimal if it matches the quantity $Y$ and a deviation of $Y$ from $a$ leads to a loss that increases quadratically with the deviation. $Y$ is a function of the random inputs $\mathbf{X}$.

In reliability applications, the loss function is typically

$$L(\mathbf{X}, a) = c(a) + \mathbb{I}[g(\mathbf{X}, a) \leq 0] c_F,$$

(3)

where $c$ is a function giving the cost of implementing the decision alternative $a$ and $c_F$ is the cost of failure. $\mathbb{I}$ is the indicator function, which results in 1 if the argument holds and 0 otherwise. The corresponding expected loss is

$$E_{\mathbf{X}} [L(\mathbf{X}, a)] = c(a) + \mathbb{E}[g(\mathbf{X}, a) \leq 0] c_F.$$  

(4)

In some cases it is possible to obtain information on $\mathbf{X}$ prior to making the decision $a$. When data $\mathbf{d}$ is available, it can be utilized to update the probability distribution of $\mathbf{X}$ through Bayesian analysis. An a-posteriori optimal decision can then be found as

$$a_{\text{opt}|d} = \arg \min_a E_{\mathbf{X}|d} [L(\mathbf{X}, a)],$$

(5)

wherein $E_{\mathbf{X}|d}$ denotes the expected value with respect to the conditional distribution of $\mathbf{X}$ given $\mathbf{d}$.

By subtracting the corresponding a-posteriori expected loss from the expected loss achieved under the a-priori optimal decision $a_{\text{opt}}$, one finds the so-called conditional value of information:

$$CVOI(\mathbf{d}) = E_{\mathbf{X}|d} [L(\mathbf{X}, a_{\text{opt}})] - E_{\mathbf{X}|d} [L(\mathbf{X}, a_{\text{opt}|d})].$$

(6)

Since $a_{\text{opt}|d}$ results in the minimum a-posteriori expected loss, it is $E_{\mathbf{X}|d} [L(\mathbf{X}, a_{\text{opt}|d})] \leq E_{\mathbf{X}|d} [L(\mathbf{X}, a_{\text{opt}})]$ and the CVOI cannot be negative.

Initially, the data $\mathbf{d}$ is not yet available. To understand the potential value of collecting the data, one evaluates the expected value of Eq. 6 over the distribution of the data. This is the (expected) value of information:

$$EVOI = E_{\mathbf{d}} [CVOI(\mathbf{d})].$$

(7)

Since the CVOI cannot be negative, also the EVOI cannot be negative.

Computation of the EVOI is non-trivial in the general case. Algorithms for efficient computation have been proposed in the literature in different application contexts [e.g., 40, 42, 44, 45].

In the limit, the decision maker is able to obtain perfect information, such that $\mathbf{X}$ becomes known with certainty. For such a clairvoyant decision maker, finding the optimal $a$ reduces to a deterministic decision problem:

$$a_{\text{opt}|x} = \arg \min_a L(\mathbf{x}, a).$$

(8)

By analogy with Eq. 7, the expected value of perfect information is

$$EVPI = E_{\mathbf{x}} [L(\mathbf{X}, a_{\text{opt}}) - L(\mathbf{X}, a_{\text{opt}|x})].$$

(9)

The EVPI is the upper bound on the EVOI, i.e., the best data collection cannot provide a higher expected value than EVPI in the considered decision context.

A special case of value of information arises when obtaining perfect information on a subset of input parameters. Consider that input random variable $X_i$ is observed to take value $x_i$. One can evaluate the conditional value of partial perfect information on $X_i$ as

$$CVPPI_{X_i}(x_i) = E_{\mathbf{X}_{-i}} [L(\mathbf{X}, a_{\text{opt}})|X_i = x_i] - E_{\mathbf{X}_{-i}} [L(\mathbf{X}, a_{\text{opt}|x_i})|X_i = x_i],$$

(10)

wherein $\mathbf{X}_{-i}$ indicates the vector of all random variables $\mathbf{X}$ except $X_i$. 

3
By evaluating the expected value of the CVPPI over the prior distribution of \( X_i \), one obtains the expected value of partial perfect information:

\[
EVPI_{X_i} = E_{X_i} \left[ CVPPi_{X_i}(X_i) \right] \\
= E_X \left[ L(X, a_{opt}) \right] - E_X \left[ L(X, a_{opt|X_i}) \right]. 
\] (11)

The EVPI is a decision-theoretic metric for factor prioritization. It describes the expected gains due to improved decision making when learning a specific input random variable.

**Remark 2.1:** When the loss function is of the quadratic form of Eq. 2, the optimal decision a-priori is \( a_{opt} = E_X[Y(X)] \). The associated expected loss is \( E_X[L(X, a_{opt})] = Var_X[Y(X)] \). Correspondingly, the expected loss associated with the optimal posterior decision is \( E_{X_{\text{opt}}}[L(X, a_{opt|x_i})|X_i = x_i] = Var_{X_{\text{opt}}}[Y(X)|X_i = x_i] \). Inserting these expressions into Eq. 11 gives

\[
EVPI_{X_i} = Var_X[Y(X)] - E_{X_i} \{ Var_{X_{\text{opt}}}[Y(X)|X_i] \} \\
= Var_{X_i} \{ E_{X_{\text{opt}}}[Y(X)|X_i] \}. 
\] (12)

Dividing this result with \( Var_X[Y(X)] \) gives the first-order variance-based sensitivity index, the Sobol’ index

Hence, the EVPI associated with a quadratic loss function is equivalent to the first-order Sobol’ index. The EVPI is also related to other sensitivity metrics, as discussed in [3, 46].

**Remark 2.2:** The definition of the EVPI can be extended to groups of input random variables \( X_v = \{X_i, i \in v\} \), with \( v \) being a subset of \( \{1, \ldots, n\} \). In the case of the quadratic loss function, the resulting \( EVPI_{X_v} \) normalized with \( Var_X[Y(X)] \) is equal to the closed Sobol’ index of \( X_v \).

The formulation of the decision alternatives and the loss function depends on the decision context. In Section 3, we derive the EVPI for cases when the analysis is performed to assess the safety of a given system or design. In Section 4 we consider the case of reliability-based optimization, in which a includes a set of optimization parameters. For both cases, we discuss efficient estimators for the EVPI. In the context of reliability sensitivity, efficiency is measured in terms of the number of additional evaluations of \( g(x) \) after the reliability is computed.

3. Safety assessment

3.1. Decision analysis

A common decision situation in engineering is associated to the safety assessment of an existing system. In its simplest form, the problem can be represented by the decision tree of Figure 1. It has two decision alternatives \( a = \{a_0, a_r\} \), corresponding to doing nothing and replacing/strengthening the system. The state of the system is binary, it either fails \( F = \{g(X) \leq 0\} \) or survives \( F^c = \{g(X) > 0\} \). The loss function is of the form given by Eq. 3.

We make the assumption that following a replacement or strengthening action \( a_r \), the system has a negligible probability of failure \(^2\) hence \( Pr(F|a_r) = 0 \). This model is equivalent to the one of [27], in which \( a_r \) is the decision to reject a component in a system design. The resulting expected loss is

\[
E_X[L(X, a)] = \begin{cases} 
  p_F c_F, & a = a_0 \\
  c_r, & a = a_r.
\end{cases}
\] (13)

where \( p_F = Pr[g(X) \leq 0] \), \( c_F \) is the cost of failure and \( c_r \) is the cost of replacement.

---

2Under a quadratic loss function, \( Var_X[Y(X)] \) is the expected value of perfect information EVPI according to Eq. 9 since the expected loss under perfect information is zero. Therefore, the first-order Sobol’ index corresponds to the EVPI normalized with the EVPI.

3While this is not the case in reality, it can be expected that the probability of failure of the updated system \( Pr(F|a_r) \) is indeed much lower than the original \( Pr(F) \).
Figure 1: Decision tree representing the safety assessment.

In this case, the optimal decision a-priori is

\[
a_{\text{opt}} = \begin{cases}  
a_0, & p_F \leq \frac{c_r}{c_F} \\
a_r, & \text{else,} \end{cases}
\]  

(14)

wherein \( \frac{c_r}{c_F} \) serves as a decision threshold.

When \( X_i \) is known, the analysis is performed with \( p_F(x_i) = \Pr[g(X) \leq 0|X_i = x_i] \), the conditional probability of failure given \( X_i = x_i \). The conditionally optimal decision is

\[
a_{\text{opt}|X_i}(x_i) = \begin{cases}  
a_0, & p_F(x_i) \leq \frac{c_r}{c_F} \\
a_r, & \text{else.} \end{cases}
\]  

(15)

If the conditional \( p_F(x_i) \) is on the same side of the threshold \( \frac{c_r}{c_F} \) as the unconditional \( p_F \), the optimal decision is not altered by the knowledge \( X_i = x_i \). In this case, the CVPPI is zero. The CVPPI is only positive, if the decision is changed. Hence it is

\[
CVPPI_{X_i}(x_i) = \begin{cases}  
|c_Fp_F(x_i) - c_r|, & a_{\text{opt}|X_i}(x_i) \neq a_{\text{opt}} \\
0, & a_{\text{opt}|X_i}(x_i) = a_{\text{opt}}. \end{cases}
\]  

(16)

The CVPPI is illustrated in Figure 2.

The resulting EVPPI is

\[
EVPPI_{X_i} = \int_{\Omega_{X_i}} |c_Fp_F(x_i) - c_r| f_{X_i}(x_i) dx_i,
\]  

(17)

wherein the integration domain \( \Omega_{X_i} \) is the set of values \( x_i \) for which the conditionally optimal decision \( a_{\text{opt}|X_i}(x_i) \) differs from the unconditional optimum \( a_{\text{opt}} \).

For example, if the optimal prior decision is \( a_{\text{opt}} = a_0 \), then the domain \( \Omega_{X_i} \) consists of all \( x_i \) for which the following inequality holds:

\[
p_F(x_i) > \frac{c_r}{c_F}.
\]  

(18)

In most cases, \( p_F(x_i) \) is a monotonic function. In these cases, a threshold value \( x_{i,\text{thres}} \) can be found at which the optimal decision changes from \( a_0 \) to \( a_r \), or reversely. It can be determined from the condition

\[
p_F(x_{i,\text{thres}}) = \frac{c_r}{c_F}.
\]  

(19)
Figure 2: Conditional probability of failure for input random variable $R$ of the example in Section 5.1, together with the conditional value of partial perfect information CVPPI. The cost of failure is $c_F = 10^8$ and the cost of replacement $c_r = 10^6$, hence the threshold is $\frac{c_r}{c_F} = 10^{-2}$. The prior probability of failure is $\Pr(F) = 7.4 \cdot 10^{-3}$, hence $a_{opt} = a_0$ and the CVPPI is non-zero only for values of $p_F(r) > 10^{-2}$.

The evaluation of the EVPPI following Eq. 17 requires the function $p_F(x_i)$. Two strategies for estimating this function are presented in the following subsections.

Remark 3.1: When evaluating the EVPPI for a group of input random variables $X_v = \{X_i, i \in v\}$, Eqs. 13 to 16 hold with $x_i$ replaced by $x_v$. The EVPPI is

$$
EVPPI_{X_v} = \int_{\Omega_{X_v}} |c_F p_F(x_v) - c_r| f_{X_v}(x_v) \, dx_v,
$$

wherein the integration domain $\Omega_{X_v}$ is the set of values for which $a_{opt|X_v}(x_v) \neq a_{opt}$.

The computational strategies for evaluating $EVPPI_{X_i}$ that we present in the following can be extended to the computation of $EVPPI_{X_v}$, but we do not discuss this further.

3.2. Estimation based on failure samples

When estimating $p_F$ with a sampling-based method, the function $p_F(x_i)$ can be estimated from samples in the failure domain [47]. To this end, we note that the conditional $p_F(x_i) = \Pr[F|X_i = x_i]$ can be found through Bayes’ rule as

$$
p_F(x_i) = \frac{f_{X_i|F}(x_i)}{f_{X_i}(x_i)} p_F,
$$

wherein $f_{X_i|F}$ is the PDF of $X_i$ conditional on a failure event.

The relationship of Eq. 21 is illustrated in Figure 3. The fact that the EVPPI is a function of the relation between $f_{X_i|F}$ and $f_{X_i}$ points to the link to distribution-based sensitivity metrics, which is discussed in [48].

If crude Monte Carlo simulation is utilized to estimate $p_F$, then the set of samples that fall into the failure domain are independent and identically distributed (iid) samples from $f_{X_i|F}$. Hence they can be utilized to obtain an estimate of $f_{X_i|F}$, e.g., by means of a kernel density estimator (KDE) as illustrated in Figure 4.

If more advanced sampling techniques are employed to estimate $p_F$, an additional step might be necessary before estimating $f_{X_i|F}$. In the case of importance sampling techniques, the resulting samples in
Figure 3: Conditional and unconditional PDF of the input random variable $R$ of the example in Section 5.1 together with the conditional failure probability $p_F(r) = \Pr(F|R=r)$.

Figure 4: Histogram of failure samples of the input random variable $M_1$ obtained as a side product from a Monte-Carlo-based reliability analysis, together with the corresponding kernel density estimate KDE. Taken from the example in Section 5.2. The number of failure samples is $n_F = 947$. Additionally, the unconditional (prior) PDF of $M_1$ is shown.
3.3. FORM-based approximation

A powerful approximation method to compute $\Pr[F]$ is the First-order Reliability Method (FORM) \cite{FORM}. FORM is based on a transformation of the limit-state function $g(X)$ to an equivalent $G(U)$ defined in terms of independent standard normal random variables $U$. This is achieved by an iso-probabilistic transform, typically the Rosenblatt \cite{Rosenblatt} or the Nataf \cite{Nataf} transform. The probability of failure is then $p_F = \Pr[G(U) \leq 0]$ and can be approximated by a linearization of $G(U)$ at the so-called most likely failure point $u^\ast$. This is the point in the failure domain $\{G(U) \leq 0\}$ with the highest probability density. Because of the rotational symmetry of the standard normal PDF, this is also the point in $\{G(U) \leq 0\}$ closest to the origin:

$$u^\ast = \arg \min_u \|u\| \quad \text{s.t. } G(u) \leq 0.$$  \hspace{1cm} (22)

$\|\cdot\|$ is the Euclidean norm.

Let $G_1$ denote the linearized limit-state function. Provided that $u^\ast$ is not the origin, it can be shown that $\Pr[G_1(U) \leq 0] = \Phi[-\beta_0]$, where $\beta_0 = \|u^\ast\|$ is the distance of the most likely failure point from the origin and $\Phi$ is the standard normal CDF \cite{CDF}. Hence the FORM approximation of the probability of failure is

$$p_F \approx \Phi[-\beta_0].$$ \hspace{1cm} (23)

The FORM $\alpha$-factors $\alpha = [\alpha_1; \ldots; \alpha_n]$ are the directional cosines of $u^\ast$ and are commonly used as a sensitivity metric \cite{alpha_factors_1} \cite{alpha_factors_2} \cite{alpha_factors_3}.

To evaluate the EVPPI, we note that the FORM estimate of the probability of failure conditional on $U_i = u_i$ is

$$p_F(u_i) = \Phi \left( \frac{-\beta_0 + \alpha_i u_i}{\sqrt{\alpha_i^2 + \alpha_{i+1}^2 + \cdots + \alpha_n^2}} \right) = \Phi \left( \frac{\alpha_i u_i - \beta_0}{\sqrt{1 - \alpha_i^2}} \right) \quad \text{for } \alpha_i \neq 0. \quad \text{For } \alpha_i = 0: \quad p_F(u_i) = \Phi[-\beta_0].$$ \hspace{1cm} (24)

In the case of independent random variables, $X_i$ is related to $U_i$ by the marginal transformation

$$u_i = \Phi^{-1}[F_{X_i}(x_i)],$$ \hspace{1cm} (25)

hence in this case it is

$$p_F(x_i) = \Phi \left( \frac{\alpha_i \Phi^{-1}[F_{X_i}(x_i)] - \beta_0}{\sqrt{1 - \alpha_i^2}} \right).$$ \hspace{1cm} (26)

Setting $p_F(x_i, \text{thres}) = \frac{c_F}{c_r}$ results in

$$x_i, \text{thres} = F_{X_i}^{-1} \left( \Phi \left( \frac{1}{\alpha_i} \left[ \sqrt{1 - \alpha_i^2} \Phi^{-1} \left( \frac{c_F}{c_r} \right) + \beta_0 \right] \right) \right).$$ \hspace{1cm} (27)

The threshold of Eq. \cite{threshold} is the boundary of the integration domain in the EVPPI calculation, Eq. \cite{EVPPI}.

Inserting Eq. \cite{EVPPI} into Eq. \cite{EVPPI} results in a FORM estimate of the EVPPI. Integration by substitution with $\Phi^{-1}[F_{X_i}(x_i)] = u_i$ gives

$$\text{EVPPI}_{X_i} = \int_{\Omega_{u_i}} [c_F p_F(u_i) - c_r |\varphi(u_i)|] \, du_i,$$ \hspace{1cm} (28)
Figure 5: EVPPI in function of the FORM sensitivity index $\alpha_i$, for varying values of $p_F$, $c_F = 1$ and $\frac{c_r}{c_F} = 10^{-3}$.

wherein $p_F(u_i)$ is given by Eq. 24, $\varphi$ is the standard normal PDF and the integration domain is

$$\Omega_{u_i} = \begin{cases} [u_{i,\text{thres}}, \infty), & p_F \leq \frac{c_r}{c_F}, \alpha_i > 0 \\ [-\infty, u_{i,\text{thres}}], & p_F \leq \frac{c_r}{c_F}, \alpha_i < 0 \\ [-\infty, u_{i,\text{thres}}], & p_F > \frac{c_r}{c_F}, \alpha_i > 0 \\ [u_{i,\text{thres}}, \infty), & p_F > \frac{c_r}{c_F}, \alpha_i < 0. \end{cases} \tag{29}$$

$u_{i,\text{thres}}$ is the transformation of the threshold $x_{i,\text{thresh}}$ to standard normal space:

$$u_{i,\text{thres}} = \frac{1}{\alpha_i} \left[ \sqrt{1 - \alpha_i^2 \Phi^{-1} \left( \frac{c_r}{c_F} \right)} + \beta_0 \right]. \tag{30}$$

In [53], we show that the integral of Eq. 28 can alternatively be evaluated as

$$EVPPI_{X_i} = |c_F \Phi_2(-\beta_0, s_i u_{i,\text{thres}}, -s_i \alpha_i) - c_r \Phi (s_i u_{i,\text{thres}})|. \tag{31}$$

$\Phi_2(x_1, x_2, r)$ is the bivariate standard normal CDF with correlation coefficient $r$ evaluated at $x_1$ and $x_2$ and $s_i$ is either $+1$ or $-1$, depending on the integration domain: $s_i = \text{sgn} \left[ (p_F - \frac{c_r}{c_F}) \alpha_i \right]$ if $(p_F - \frac{c_r}{c_F}) \alpha_i \neq 0$, otherwise $s_i = -1$.

Remark 3.2: The FORM approximation of the EVPPI is identical for $+\alpha_i$ and $-\alpha_i$, hence it is also possible to utilize $|\alpha_i|$ in the above expressions.

3.4. Relation of EVPPI to the FORM sensitivity index

Equations 28-30 show that the FORM approximation of the EVPPI does not depend on the marginal distribution of $X_i$. It is a function only of $\beta_0$, $\alpha_i$ and $\frac{c_r}{c_F}$. Hence, one can directly obtain the EVPPI corresponding to a FORM sensitivity $\alpha_i$ without any calls of the $g$-function.

Figure 5 shows the EVPPI in function of $|\alpha_i|$ for different values of $p_F$, with cost ratio $\frac{c_r}{c_F} = 10^{-3}$; Figure 6 shows the same for cost ratio $\frac{c_r}{c_F} = 10^{-4}$. The EVPPI is highest when $p_F$ is close to $\frac{c_r}{c_F}$. This is to be expected, because in these cases the observation $X_i = x_i$ is most likely to lead to a change in the optimal decision.

The EVPPI is increasing with increasing value of $|\alpha|$. Since $\beta_0$ and $\frac{c_r}{c_F}$ are fixed for a given reliability problem, the ordering according to FORM approximation of the EVPPI is the same as the ordering according to the absolute values of the FORM sensitivities $|\alpha_i|$. This holds for reliability problems with independent input random variables when the decision problem is as in Figure 1.
The key benefit of using the EVPPI metric rather than the FORM α-factors is that it accounts for the effect of β0 and c_F on the decision sensitivity. For example, consider the case c_F = 10^{-3} shown in Figure 5. One can observe that for p_F = 10^{-3} an input random variable with |α| = 0.8 has an EVPPI that is double the one associated with |α| = 0.35. In contrast, for p_F = 10^{-2} the EVPPI corresponding to |α| = 0.8 is 34 times larger than the EVPPI corresponding to |α| = 0.35. Hence the ratio between |α| (or α^2) values is not a good indicator for how much more important one input random variable is relative to another in the specific decision context. But the ratio between EVPPI values is; it is admissible to make a statement like “learning X_i is \frac{EVPPI_{X_i}}{EVPPI_{X_j}} more valuable than learning X_j.”

3.5. Normalization

It can be preferable to express the sensitivity by an index that is normalized to lie between 0 and 1. A trivial normalization can be achieved by normalizing the EVPPI by the sum of the EVPPI of all input random variables:

\[
\text{Normalized } EVPPI_{X_i} = \frac{EVPPI_{X_i}}{\sum_{j=1}^{n} EVPPI_{X_j}}.
\]  

(32)

The advantage of this normalization is that the indices sum up to one. Therefore, this normalization is suitable for comparison with other sensitivity indices that have this property or are normalized accordingly.

A natural normalization is obtained by evaluating the EVPPI relative to the expected value of perfect information (EVPI) of Eq. 9. The EVPI is the value associated with having full information on all inputs. We refer to this normalized index as the relative EVPPI:

\[
\text{Relative } EVPPI_{X_i} = \frac{EVPPI_{X_i}}{EVPI}.
\]  

(33)

For the safety assessment decision, the EVPI is

\[
EVPI = \begin{cases} 
  p_F (c_F - c_r), & p_F \leq \frac{c_r}{c_F} \\
  c_r (1 - p_F), & \text{else,} 
\end{cases}
\]  

(34)

and is readily available.

Because the EVPI is an upper bound on the EVPPI, the relative EVPPI is bounded by zero and one. However, the relative EVPPIs do not sum up to one. Furthermore, their sum can be larger than one, because the EVPPI for groups of input random variables can be lower than the sum of the EVPPIs of individual input random variables [54].

Remark 3.3: As discussed in Section 2 the Sobol’ index is a relative EVPPI for the case of the quadratic loss function.
4. Reliability-based design

4.1. Decision analysis

In reliability-based design, the decision is typically not binary as in the case of the safety assessment. Instead, one or more design parameters \( a \) can be selected from an (often continuous or continuous-discrete) domain. The function \( c_d(a) \) describes the design cost, which is here assumed to be deterministic, i.e., it does not depend on \( X \). As in Section 3, the cost of a failure is \( c_F \). Hence the loss function is

\[
L(X, a) = c_d(a) + I[g(X, a) \leq 0]c_F. \tag{35}
\]

The optimal decision a-priori is

\[
a_{\text{opt}} = \arg \min_a c_d(a) + p_F(a)c_F. \tag{36}
\]

The a-posteriori optimal decision given \( X_i = x_i \) is

\[
a_{\text{opt}|X_i}(x_i) = \arg \min_a c_d(a) + p_F(x_i, a)c_F. \tag{37}
\]

\( p_F(x_i, a) = \Pr\{g(X, a) \leq 0 | X_i = x_i\} \) is the conditional probability of failure given \( X_i = x_i \) for a design \( a \).

Following Eq. 11 the EVPPI is

\[
EVPPI_{X_i} = c_d(a_{\text{opt}}) + p_F(a_{\text{opt}})c_F - E_{X_i}\{c_d[a_{\text{opt}|X_i}(X_i)] + p_F[X_i, a_{\text{opt}|X_i}(X_i)]c_F\}. \tag{38}
\]

This computation requires that \( p_F(x_i, a) \) is available in a manner that facilitates the solution of Eq. 37. This optimization problem has to be solved many times for varying values of \( X_i \), hence the evaluation of \( p_F(x_i, a) \) must be efficient and without noise. In general, \( p_F(x_i, a) \) will not be available in analytical form. We present a computational strategy to address this challenge in the following. In Section 4.3 we additionally discuss a FORM approximation.

4.2. Discrete or discretized design choices

In the case where the domain of \( a \) is discrete, or can be discretized, one can solve many reliability problems, one for each design choice \( \{a_1, \ldots, a_m\} \). For a given \( a_j \), the conditional \( p_F(x_i, a_j) \) can be determined with the strategies presented in Sections 3.2 and 3.3. The solution of Eq. 37 then reduces to selecting the optimal \( a_j \) from the discrete set \( \{a_1, \ldots, a_m\} \).

When the domain of \( a \) is continuous, the consideration of only a discrete set of design choices \( \{a_1, \ldots, a_m\} \) leads to an approximation error. The difference between the exact and approximated expected a-posteriori loss \( E_{X_i}[L(X, a_{\text{opt}|X_i}(X_i)) = c_d[a_{\text{opt}|X_i}(x_i)] + p_F[x_i, a_{\text{opt}|X_i}(x_i)]c_F \), is illustrated in Figure 7. The expected loss under the a-posteriori optimal choice \( a_{\text{opt}|X_i} \) is overestimated for all values of \( X_i \), except for those where the discrete choices \( a_j \) coincide with the optimal choice in the continuous domain of \( a \).

To limit the error from this approximation, the discrete design choices should cover all values of \( a \) that might potentially be optimal. The spacing between \( a_j \)'s should represent a good trade-off between accuracy and computational cost. These choices are investigated numerically in Section 5.1.5.

4.3. A remark on the FORM approximation

It is possible to derive a FORM approximation of the EVPPI also in the design decision case. Such an approximation is described in the Annex of the paper, assuming a linear design LSF.

The EVPPI for different values of the squared FORM \( \alpha \)-values relative to the EVPPI of \( \alpha^2 = 1 \) are shown in Figure 8. These results are independent of the cost parameters \( c_d \) and \( c_F \) introduced in the Annex.

Based on the results of the Annex and Figure 8, one can conclude that for any design LSF with independent inputs, whose reliability is well approximated by FORM, one should expect the relative EVPPI of the design decision case to give similar importance measures than the FORM \( \alpha \)-factors.
Figure 7: Expected loss $E[L(X,a)|S=s]$ conditional on input random variable $S$, from the example in Section 5.1. The black line is the posterior expected loss under the a-posteriori optimal choice, $E[L(X,a_{opt}|s)|S=s]$. In contrast, the colored lines represent the posterior expected loss for fixed design choices, $E[L(X,a_i)|S=s]$. The minimum of these would apply for a specific value of $s$. The black and the colored lines meet at the value of $s$ at which the fixed choice $a_i$ is identical to the a-posteriori optimal choice $a_{opt|s}$.

Figure 8: EVPPI for the reliability-based design case evaluated from the FORM approximation, normalized with the EVPPI at $\alpha^2 = 1$. The three curves correspond to a reliability index $\beta_0 = 3$ (blue), 4 (black) and 5 (yellow). Details are provided in the Annex.
### 4.4. Normalization

By analogy with Section 3.5, the normalized EVPPI and the relative EVPPI can also be defined for the reliability-based design. While computation of the former is trivial, the computation of the EVPI necessary for determining the relative EVPPI generally requires additional model runs. In most cases, this additional effort is not justified.

### 5. Numerical investigations

#### 5.1. Example 1: Component reliability

##### 5.1.1. VOI sensitivity for safety assessment

We consider a system with resistance $R$ and load $S$ and two model uncertainties $X_R$ and $X_S$. The system fails when $X_S S$ exceeds $X_R R$:

$$F = \{ X_R R \leq X_S S \}. \quad (39)$$

A limit-state function describing this failure event is

$$g(\textbf{X}) = \ln X_R + \ln R - \ln X_S - \ln S. \quad (40)$$

The probabilistic model of $\textbf{X} = [R; S; X_R; X_S]$ is summarized in Table 1.

Because all random variables are modelled as lognormal, the probability of failure can be computed analytically. It is

$$p_F = \Phi \left( \frac{-\ln X_R + \mu \ln R - \mu \ln X_S - \mu \ln S}{\sqrt{\sum_{i,j} C_{ij}}} \right), \quad (41)$$

wherein $\textbf{C}$ is the covariance matrix of $\ln \textbf{X}$.

The conditional probabilities of failure $p_F(x_i)$ are evaluated by first computing (analytically) the conditional moments of $\ln X_i$ given $X_i = x_i$. $p_F(x_i)$ is then obtained by analogy with Eq. 41.

The EVPPI is evaluated (a) assuming independent input random variables and (b) assuming that the input random variables follow a Gaussian copula, with correlation matrix of $\textbf{X}$ equal to

$$C_{XX} = \begin{bmatrix} 1 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0.5 & 0 & 1 & 0.5 \\ 0 & 0.5 & 0.5 & 1 \end{bmatrix}.$$}

The probability of failure in the case of independent input random variables is $\text{Pr}(F) = 7.4 \cdot 10^{-3}$ and in the case of dependent input random variables it is $\text{Pr}(F) = 1.7 \cdot 10^{-2}$.

Cost values are $c_F = 10^5$ and either $c_r = 10^5$ or $c_r = 10^6$, corresponding to cost ratios of either $\frac{c_r}{c_F} = 10^{-3}$ or $\frac{c_r}{c_F} = 10^{-2}$.

Figure 3 shows the relation between the prior PDF $f_R$, the conditional PDF $f_{RF}$ and the resulting conditional failure probability $p_F(r)$ for the input random variable $R$ (independent case). Figure 2 shows the resulting $\text{CVPPI}(r)$ for $\frac{c_r}{c_F} = 10^{-2}$. 

### Table 1: Probabilistic model of example 1.

| Parameter          | Distribution | Mean | c.o.v. |
|--------------------|--------------|------|--------|
| Resistance $R$     | lognormal    | 100  | 0.2    |
| Load $S$           | lognormal    | 40   | 0.25   |
| Model uncertainty $X_R$ | lognormal | 1    | 0.1    |
| Model uncertainty $X_S$ | lognormal | 1    | 0.2    |

13
Table 2: Optimal designs, together with the associated failure probabilities.

| $a_{opt}$ | $p_F(a_{opt})$ | $c_d(a) = 10^5a$ | $c_d(a) = 10^6a$ |
|-----------|----------------|------------------|------------------|
| independent $X$ | $1.57$ | $1.87$ | $1.23$ |
| dependent $X$ | $2.7 \cdot 10^{-4}$ | $1.23$ | $1.41$ |

Table 3: Normalized sensitivity metrics for the component example with independent input random variables.

| Normalized EVPI FORM | Normalized EVPI FORM |
|----------------------|----------------------|
| safety $\frac{\partial}{\partial \alpha} = 10^{-3}$ | safety $\frac{\partial}{\partial \alpha} = 10^{-2}$ | design $\frac{\partial}{\partial \alpha} = 10^{-3}$ | design $\frac{\partial}{\partial \alpha} = 10^{-2}$ |
| Resistance $R$ | $25\%$ | $27\%$ | $26\%$ | $26\%$ |
| Load $S$ | $49\%$ | $35\%$ | $41\%$ | $41\%$ |
| Model uncertainty $X_R$ | $0.5\%$ | $10\%$ | $7\%$ | $6\%$ |
| Model uncertainty $X_S$ | $25\%$ | $27\%$ | $26\%$ | $26\%$ |

5.1.2. VoI sensitivity for reliability-based design

For design optimization, we consider a modified version of the limit-state function,

$$g(X) = aX_R R - X_S S,$$

in which $a$ is the design parameter affecting the overall resistance of the system.

The probability of failure in function of $a$ is

$$p_F(a) = \Phi \left[ -\ln a + \mu_{\ln X_R} + \mu_{\ln R} - \mu_{\ln X_S} - \mu_{\ln S} \right].$$

For simplicity, we assume that the design cost is a linear function of $a$:

$$c_d(a) = c_0 a.$$

Two cases are evaluated, with cost factor equal to $c_0 = 10^5$ and $c_0 = 10^6$.

The optimal design a-priori is found by Eq. 36 and reported in Table 2, together with the associated failure probability.

5.1.3. Results and discussion

Tables 3 and 4 summarize the resulting decision sensitivities for selected cost models. In addition, FORM sensitivities are reported in terms of the squared $\alpha$-factors or the $\gamma^2$, which are a generalization of the $\alpha$-factors for dependent inputs [14, 15]. The EVPPI are stated as normalized values to facilitate a comparison with the squared $\alpha$-factors.

As expected for this linear Gaussian case, the ranking coincides for all EVPPI cases and the FORM sensitivities. Note that in the independent case, this equivalence follows from Sections 3.3 and 4.3.

Table 4: Normalized sensitivity metrics for the component example with dependent input random variables.

| Normalized EVPI FORM | Normalized EVPI FORM |
|----------------------|----------------------|
| safety $\frac{\partial}{\partial \alpha} = 10^{-3}$ | safety $\frac{\partial}{\partial \alpha} = 10^{-2}$ | design $\frac{\partial}{\partial \alpha} = 10^{-3}$ | design $\frac{\partial}{\partial \alpha} = 10^{-2}$ |
| Resistance $R$ | $15\%$ | $26\%$ | $23\%$ | $23\%$ | $26\%$ |
| Load $S$ | $61\%$ | $41\%$ | $46\%$ | $46\%$ | $41\%$ |
| Model uncertainty $X_R$ | $0.0\%$ | $4\%$ | $4\%$ | $4\%$ | $7\%$ |
| Model uncertainty $X_S$ | $24\%$ | $29\%$ | $28\%$ | $28\%$ | $26\%$ |
However, for the safety decision, the relative contribution varies and is notably different for cost ratio \( \frac{c_R}{c_F} = 10^{-3} \). In this case, the EVPPI of input \( X_R \) is negligible, which is due to the fact that at this cost ratio, it is very unlikely that learning about \( X_R \) will change the a-priori optimal decision.

The design EVPPI results of the independent case are similar to the squared \( \alpha \)-factors, which confirms the conclusions in Section 4.3. Generally, the results for the EVPPI sensitivity of the reliability-based design case are independent of the ratio \( \frac{c_F}{\delta} \) for this linear Gaussian model. This is in agreement with the findings reported in the Annex.

Figure 9 plots the EVPPI of the safety decision case in function of \( \frac{c_R}{c_F} \) for the case with dependent inputs. (The plot for independent inputs looks similar.) It compares the absolute value of the EVPPI with the relative and the normalized EVPPI. Both the absolute and the relative EVPPI are highest with a decision threshold \( \frac{c_R}{c_F} \) close to the probability of failure \( p_F \). This is not surprising, since in this case it is more likely that learning an input parameter will change the optimal decision. The normalized EVPPI shows that when \( \frac{c_R}{c_F} \) is close to \( p_F \), the differences between the EVPPIs are smaller. Near the threshold, even learning the less important input random variables still has considerable value. The further \( \frac{c_R}{c_F} \) deviates from \( p_F \), the larger the differences between EVPPIs. In the extreme, the normalized EVPPI of the most important input random variable will approach 1.

5.1.4. Monte Carlo approximation for the safety decision case

To investigate the accuracy of the Monte Carlo approach of Section 3.2, we determine the EVPPI with \( 10^2 \) and \( 10^3 \) failure samples. This is the typical range of the number of samples in the failure domain one would expect from a sampling-based reliability evaluation.

The results are summarized in Table 5. Each analysis is repeated 100 times, to obtain an approximate mean and standard deviation of the MCS estimate. The results show that for this example the accuracy of the estimate is sufficient for as low as \( n_F = 100 \) independent failure samples.

5.1.5. Discrete approximation for the design decision case

We investigate the discrete approximation of design choices following Section 4.2. Results are shown here for the setting \( c_d = 10^5 \) and independent input random variables.

Discrete design parameter choices between \( a_1 = 0.5 \) and \( a_m = 2.0 \) are considered, with step size \( \Delta a = \frac{1.5}{m-1} \). The approximation of the posterior expected loss in case of \( m = 4 \) is shown in Figure 7.

Figure 10 shows the convergence of estimated EVPPI to the exact values with increasing \( m \). While an accurate evaluation of the absolute EVPPI here requires \( m \geq 6 \) discrete design choices, the normalized EVPPI is already evaluated reasonably well with \( m \) as low as 3, see Table 6.

15
Table 5: EVPPI for the safety-assessment case with $c_r = 10^{-2}$, $c_F = 10^6$ and independent input random variables. Exact value and sample-based value with $n_F$ failure samples. For the latter, the mean value and coefficient of variation (in parenthesis) are evaluated by 100 repetitions of the analysis.

|             | Exact | $n_F = 10^3$ | $n_F = 10^4$ |
|-------------|-------|--------------|--------------|
| Resistance R | 349   | 344(3.2%)    | 337(7.1%)    |
| Load S      | 454   | 443(2.3%)    | 443(6.0%)    |
| Model uncertainty $X_R$ | 131   | 131(8.3%)    | 127(22.4%)   |
| Model uncertainty $X_S$ | 349   | 344(3.2%)    | 339(7.9%)    |

Table 6: Normalized EVPPI for the reliability-based design case with $c_\delta = 10^5$ and independent input random variables.

|             | Exact | $m=3$ | $m=4$ | $m=6$ |
|-------------|-------|-------|-------|-------|
| Resistance R | 26%   | 27%   | 26%   | 26%   |
| Load S      | 41%   | 36%   | 44%   | 39%   |
| Model uncertainty $X_R$ | 7%    | 9%    | 5%    | 8%    |
| Model uncertainty $X_S$ | 26%   | 27%   | 26%   | 26%   |

Figure 10: Convergence of the EVPPI with number of discrete values $m$ of the design choice. (EVPPI evaluated for the reliability-based design case with $c_\delta = 10^5$ and independent input random variables.)
Table 7: Probabilistic model of example 2, including correlation matrix $R_{XX}$.

| Parameter                  | Distribution | Mean | c.o.v. | $R_{XX}$ |
|----------------------------|--------------|------|--------|----------|
| Bending moment $M_1$ [kNm]| normal       | 250  | 0.3    | $M_1$    |
| Bending moment $M_2$ [kNm]| normal       | 125  | 0.3    | $M_2$    |
| Axial force $P$ [kN]      | Gumbel       | 2500 | 0.2    | $P$      |
| Yield strength $Y$ [N/mm$^2$] | Weibull | 40   | 0.1    | $Y$      |

Table 8: Normalized sensitivity indices for example 2. FORM and GRIM results are taken from [15] and are provided for comparison.

| Parameter | Nor. EVPI ($c_F = 10^{-3}$) | Nor. EVPI ($c_F = 10^{-2}$) | Nor. EVPI ($c_F = 10^{-1}$) | FORM $\alpha^2$ | GRIM $\gamma^2$ |
|-----------|-----------------------------|-----------------------------|-----------------------------|------------------|------------------|
| $M_1$     | 21%                         | 22%                         | 9%                          | 7%               | 8%               |
| $M_2$     | 23%                         | 23%                         | 9%                          | 7%               | 6%               |
| $P$       | 8%                          | 24%                         | 28%                         | 20%              | 25%              |
| $Y$       | 48%                         | 31%                         | 54%                         | 65%              | 61%              |

5.2. Example 2: Non-linear limit state

We consider a short column subjected to biaxial bending moments $M_1$ and $M_2$ and axial force $P$ that was previously investigated in [14, 15]. The failure is defined by the limit-state function:

$$g(X) = 1 - \frac{M_1}{s_1Y} - \frac{M_2}{s_2Y} - \left(\frac{P}{AY}\right)^2.$$ (45)

The random variables are $X = [M_1; M_2; P; Y]$, with $Y$ being the yield strength of the material. Their joint probability distribution is a Gaussian copula model as in [15] with the parameters given in Table 7. The deterministic parameters are the flexural moduli of the plastic column section $s_1 = 0.03m$ and $s_2 = 0.015m$ and the column cross section $A = 0.190m^2$.

The reliability is evaluated with crude Monte Carlo with $10^6$ samples. The MC estimate of the probability of failure is $0.0094$ with 95% credible interval $[0.0092, 0.0096]$.

We consider only the safety assessment case. The EVPPI is determined based on the MC samples in the failure domain, following Section 3.2. Figure 4 exemplary shows the KDE fit of the conditional distribution of $M_1$ given failure.

Different cost ratios $c_F$ are considered to reflect different decision situations. The results are summarized in Table 8 together with results from the sensitivity metrics evaluated in [15].

The resulting EVPPI indicate that the decision sensitivities can vary strongly among different decision situations. For the situation $c_F = 10^{-1}$, the EVPPI is more or less comparable to the sensitivities obtained with the indices from FORM and GRIM (Generalized Reliability Importance Measures). This corresponds to the case where the cost of repair is large and hence the a-priori decision is not to repair. $\frac{c_F}{c_P} = 10^{-2}$ is very close to the a-priori probability of failure of 0.0094. In this case, learning the value of any of the four inputs is likely to change the decision from not repairing to repairing, which explains the more uniform sensitivity values. In the case $\frac{c_F}{c_P} = 10^{-3}$, the EVPPI leads to a different ranking. In this case, the prior decision is to repair. The probability of changing this decision depends on the other tails of the distributions than in cases $\frac{c_F}{c_P} = 10^{-1}$ and $\frac{c_F}{c_P} = 10^{-2}$. This explains the differences between these two decision situations, and also highlights that the EVPPI does account for the full distribution.

5.3. Example 3: Foundation reliability

We consider a monopile foundation of an offshore wind turbine in stiff, plastic soil (Figure 11). The soil-structure interaction is described by a nonlinear finite element model. Details regarding the engineering
Figure 11: Wind turbine monopile foundation [55].

The deterministic parameters of the monopile foundation are its depth \( L = 30 \) m, diameter \( d = 6 \) m, wall thickness \( t = 0.07 \) m, Poisson ratio \( \nu = 0.3 \) and Young’s modulus \( E = 2.1 \times 10^5 \) MPa. The uncertain inputs comprise the Gumbel-distributed lateral load \( H \) and its distribution parameters \( a_H \) and \( b_H \) as well as the undrained shear strength \( s \) and its hyperparameters \( s_0 \) and \( s_1 \). \( s \) is modelled as a 1-D, non-stationary, log-normal random field with a linear drift along the \( z \)-coordinate. It is discretized using the midpoint method [54] and an 82-dimensional standard-normal random vector \( \xi \). Its hyperparameters are the ground-level shear strength \( s_0 \) (at \( z = 0 \)) and the mean shear strength slope \( s_1 \). The latter two can be thought of as representing inter-site variability while \( \xi \) models intra-site variability. The probabilistic inputs are summarized in Table 9. A detailed description of the probabilistic model setup is provided in [22, 57].

| Input | Distribution  | Mean \( \mu \) | CoV \( \delta \) |
|-------|---------------|-----------------|-----------------|
| \( \xi \) | Standard-Normal | 0 | n.d. (\( \Sigma_{\xi \xi} = I_{82 \times 82} \)) |
| \( s_0 \) [kPa] | Log-Normal | 33.7094 | 0.3692 |
| \( s_1 \) [kPa] | Log-Normal | 0.7274 | 0.8019 |
| \( H \) [kN] | Gumbel | \( \mu_{H|a_H, b_H} \) | \( \delta_{H|a_H, b_H} \) |
| \( a_H \) [kN] | Log-Normal | 2274.97 | 0.2 |
| \( b_H \) [kN] | Log-Normal | 225.02 | 0.2 |

Failure occurs when the maximum stress in the foundation, \( \sigma_{\max}(X) \) exceeds \( \sigma_{\text{crit}} = 100 \) MPa. The corresponding limit-state function is

\[
g(X) = \sigma_{\text{crit}} - \sigma_{\max}(X). \tag{46}
\]

The analysis is performed with subset simulation with \( 10^4 \) samples per level [58] [59]. To evaluate the coefficient of variation of the subset simulation estimate, we repeat the computation 100 times. The a-priori
conditional failure probabilities associated with all variables but $\xi$. The coefficients of variation of the estimate (first position), absolute estimated value (second position). The coefficients of variation of the safety cases are estimated based on 100 repeated runs of subset simulation.

|          | safety $\frac{\sigma_r}{c_F}$ = 10^{-4} | safety $\frac{\sigma_r}{c_F}$ = 10^{-3} | design linear $c_d(a)$ | design quadratic $c_d(a)$ |
|----------|------------------------------------------|------------------------------------------|------------------------|------------------------|
| $s_0$    | 0\% (1.6; 4.7)                           | 2\% (0.84; 542)                          | 8\% (-; 4971)          | 8\% (-; 2790)          | 0.58\%               |
| $s_1$    | 0\% (1.81; 2.5)                          | 34\% (0.23; 7414)                       | 19\% (-; 11434)        | 14\% (-; 5084)         | 4.59\%               |
| $a_H$    | 36\% (0.63; 678)                         | 36\% (0.16; 8019)                       | 41\% (-; 24351)        | 41\% (-; 14542)        | 22.70\%              |
| $b_H$    | 63\% (0.31; 1176)                        | 28\% (0.26; 6071)                       | 32\% (-; 18792)        | 36\% (-; 12662)        | 70.39\%              |

Table 10: Normalized sensitivity measures for the monopile. In brackets: coefficient of variation of the estimate (first position), absolute estimated value (second position). The coefficients of variation of the safety cases are estimated based on 100 repeated runs of subset simulation.

| $H$ | sensitivity of the latter can be estimated, but is not meaningful, since it is not possible to reduce uncertainty on $H$ directly. It is only possible to reduce the uncertainty of its distribution parameters, $a_H$ and $b_H$.

The conditional density in the numerator of Eq. 21 is estimated with the 104 samples that are produced in the final level of subset simulation. Note that these samples are dependent, because of the MCMC algorithm that is used to propagate samples from one level to the next in subset simulation. The results are reported in Table 10. In addition to the normalized EVPPI, also the reliability-oriented Sobol’ indices of $\log p_F$ introduced in [22] are presented for comparison.

In the case $\frac{\sigma_r}{c_F} = 10^{-3}$, the EVPPI-based ranking is identical to the one of obtained with the reliability-oriented Sobol’ indices. In the case $\frac{\sigma_r}{c_F} = 10^{-2}$, the EVPPI-based ranking differs. In particular, $s_1$ becomes considerably more relevant as the repair cost is increased and $b_H$ becomes less important.

5.3.2. VoI sensitivity for reliability-based design

We consider the monopile diameter $d \in \mathbb{D}$ as the design parameter, with the feasible domain $\mathbb{D} = [5, 7]$ m. The cost of failure is $c_F = 10^8$ and the design cost is modeled with either a linear cost model $c_d(d) = 10^5 \cdot d$ or a quadratic cost model $c_d(d) = (5 + 2 \cdot ((d - 5)/2)^2) \cdot 10^5$, such that for both models, $5 \cdot 10^5 \leq c_d \leq 7 \cdot 10^5$ on $\mathbb{D}$. The two models are plotted in Figure 12. We discretize the feasible domain for the design parameter in $m = 201$ steps of $\Delta d = 0.01$m as $\{5.00, 5.01, 5.02, \ldots, 7.00\}$m and run 201 subset simulations to obtain $10^4$ failure samples associated with each discrete design. Based on these failure samples, we compute the conditional failure probabilities associated with all variables but $\xi$ and $H$ according to Eq. 21 for each
discrete design. The EVPPI results for the reliability-based design case are summarized in Table 10. The linear cost model produces more expensive designs on average thus yielding larger absolute EVPPI than the quadratic model. The difference in normalized EVPPI (percentages) between the two cost models is negligible.

6. Discussion

Many sensitivity metrics exist in the literature, also in the context of reliability assessment. The interpretation of these metrics is often difficult, even for experts [60]. In this sense, the decision-theoretic sensitivity metrics have a clear advantage: They have a straightforward interpretation in the context of a specific decision situation. This makes them applicable to all models and applications, including system reliability problems and problems with dependent inputs.

In this paper, we have shown how the expected value of partial perfect information (EVPPI) can be evaluated by post-processing the results of a reliability analysis. For the safety decision case, no additional model evaluations are necessary, hence the computation of EVPPI is cheap. For the design decision case, additional model evaluations are necessary, so the metric might not be convenient when models are computationally expensive. Even if not sufficient computational resources are available to evaluate the EVPPI for the design decision, the consideration of the decision-theoretic basis of the EVPPI metric can nevertheless be helpful for interpreting other sensitivity metrics.

If suitable surrogate models exist, efficient computation of the EVPPI in the design decision case might be possible. We have experimented with such a surrogate model approach, based on [22], and have applied it to the example of Section 5.3. However, we have found that the computation of the EVPPI requires a surrogate with high accuracy over a wide range of the input random variable’s outcome space. Standard surrogates used for reliability analysis do not have this property. The development of suitable surrogates is left for future research.

Our paper focuses on computational strategies for evaluating the EVPPI and provides only limited guidance on how to select the decision case and the associated cost/threshold parameters in specific applications. In some projects the analyst faces a situation closely resembling the decision cases presented here, in other projects less so. In the following, we discuss two applications with different decision situations, to indicate how the presented decision cases can nevertheless be utilized as a proxy if the cost/threshold parameters are chosen appropriately.

Firstly, in the analysis presented here we have assumed that the decision is taken based on minimizing expected loss. However, regulatory constraints often result in a decision that is based on a reliability criterion in the form of a safety threshold \( p_T^F \). In this case, one aims at minimizing the expected life-cycle costs under the constraint \( p_F \leq p_T^F \). This case is also considered in [28], who suggest to find a utility function such that the optimal decision is repair \((a_r)\) if \( \Pr(F) > p_T^F \) and do nothing \((a_0)\) otherwise. In our framework, this translates to setting \( \frac{c_r}{c_F} = p_T^F \) and then applying the methods of Section 3.

If the costs \( c_r \) and \( c_F \) are known, setting \( \frac{c_r}{c_F} = p_T^F \) might not reflect well the decision situation. Instead, one could include the constraint \( p_T^F \) into the decision analysis. This might, however, have the effect that the EVPPI can become negative, e.g., if the a-priori probability of failure is slightly below the acceptable \( p_T^F \). In this case, the decision maker might prefer not to gain additional information, as that might change the probability of failure estimate to an unacceptable level and incur the repair costs \( c_r \). Such information avoidance has been discussed in the literature [61]. For the purpose of sensitivity analysis, we suggest to avoid the issue, by either ignoring the safety criterion \( p_T^F \) in the EVPPI calculation, or by setting \( \frac{c_r}{c_F} = p_T^F \).

Secondly, scientific analyses are often undertaken without a specific decision in mind [46]. Or the analyst might find it difficult to specify a loss function for other reasons. In these cases, an EVPPI based on the safety decision case might be computed, with varying threshold \( \frac{c_r}{c_F} \), as depicted in Figure 9. Reporting sensitivities for varying \( \frac{c_r}{c_F} \) has the advantage that the dependence of the sensitivity on the decision context is clearly communicated.
7. Conclusion

We demonstrate that the expected value of partial perfect information (EVPPI) is an informative sensitivity metric for factor prioritization in reliability assessments, which – in contrast to other sensitivity metrics – is straightforward to interpret. It is a measure of decision sensitivity and measures how learning an input random variable improves the decision taken based on the reliability assessment. With the computational strategies proposed in this paper, the EVPPI of the safety decision case can be evaluated following a reliability analysis without any additional model runs.

Acknowledgment

We acknowledge support by the German Research Foundation (DFG) through Grants STR 1140/11-1 and PA 2901/1-1.

References

[1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis: The Primer, John Wiley & Sons, 2008.
[2] B. Iooss, P. Lemaitre, A review on global sensitivity analysis methods, in: C. Meloni, G. Dellino (Eds.), Uncertainty 649 management in Simulation-Optimization of Complex Systems: Algorithms and Applications, Springer, 2015.
[3] E. Borgonovo, E. Plischke, Sensitivity analysis: A review of recent advances, European Journal of Operational Research 248 (3) (2016) 869–887.
[4] I. M. Sobol, Sensitivity estimates for nonlinear mathematical models, Mathematical Modelling and Computational Experiment 1 (1993) 407–414.
[5] M. J. Jansen, Analysis of variance designs for model output, Computer Physics Communications 117 (1999) 35–43.
[6] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communications 181 (2010) 259–270.
[7] M. H. Chun, S. J. Han, N. I. Tak, An uncertainty importance measure using a distance metric for the change in a cumulative distribution function, Reliability Engineering and System Safety 70 (2000) 313–321.
[8] E. Borgonovo, A new uncertainty importance measure, Reliability Engineering and System Safety 92 (2007) 771–784.
[9] E. Borgonovo, B. Iooss, Moment-independent and reliability-based importance measures, in: R. Ghanem et al. (Ed.), Handbook of Uncertainty Quantification, Springer, 2016.
[10] M. Hohenbichler, R. Rackwitz, Sensitivity and importance measures in structural reliability, Civil Engineering Systems 3 (4) (1986) 203–209.
[11] H. O. Madsen, Omission sensitivity factors, Structural Safety 5 (1) (1988) 35–45.
[12] P. Bjerager, S. Krenk, Parametric Sensitivity in First Order Reliability Theory, Journal of Engineering Mechanics 115 (7) (1989) 1577–1582.
[13] O. Ditlevsen, H. O. Madsen, Structural reliability methods, Wiley, 1996.
[14] A. Der Kiureghian, First- and second-order reliability methods. Chapter 14, in Engineering design reliability handbook, edited by E. Nikolaidis et al., CRC Press, Boca Raton, FL, 2005.
[15] T. Kim, J. Song, Generalized Reliability Importance Measure (GRIM) using Gaussian mixture, Reliability Engineering and System Safety 173 (2018) 105–115.
[16] V. Chabridon, M. Baiesdent, J.-M. Bourinet, J. Morio, N. Gayton, Reliability-based sensitivity estimators of rare event probability in the presence of distribution parameter uncertainty, Reliability Engineering & System Safety 178 (2018) 164–178.
[17] G. Sarazin, J. Morio, A. Lagroux, M. Baiesdent, L. Brevault, Reliability-oriented sensitivity analysis in presence of data-driven epistemic uncertainty, Reliability Engineering & System Safety 215 (2021) 107733.
[18] I. Papaioannou, K. Breitung, D. Straub, Reliability sensitivity estimation with sequential importance sampling, Structural Safety 75 (2018) 24–34.
[19] L. J. Cui, Z. Z. Lu, X. P. Zhao, Moment-independent importance measure of basic random variable and its probability density evolution solution, Science China Technological Sciences 53 (4) (2010) 1138–1145.
[20] L. Li, Z. Lu, J. Feng, B. Wang, Moment-independent importance measure of basic variable and its state dependent parameter solution, Structural Safety 38 (2012) 40–47.
[21] P. Wei, Z. Lu, W. Hao, J. Feng, B. Wang, Efficient sampling methods for global reliability sensitivity analysis, Computer Physics Communications 185 (8) (2012) 1728–1743.
[22] M. Ehre, I. Papaioannou, D. Straub, A framework for global reliability sensitivity analysis in the presence of multi-uncertainty, Reliability Engineering and System Safety 195 (2020) 106726.
[23] S. Kucherenko, S. Song, L. Wang, Quantile based global sensitivity measures, Reliability Engineering and System Safety 185 (2019) 35–48.
[24] J. C. Felli, G. B. Hazen, Sensitivity Analysis and the Expected Value of Perfect Information, Medical Decision Making 18 (1) (1998) 95–109.
[59] I. Papaioannou, W. Betz, K. Zwirglmaier, D. Straub, Mcmc algorithms for subset simulation, Probabilistic Engineering Mechanics 41 (2015) 89–103.
[60] A. Saltelli, K. Aleksankina, W. Becker, P. Fennell, F. Ferretti, N. Holst, S. Li, Q. Wu, Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices, Environmental Modelling and Software 114 (2019) 29–39.
[61] M. Pozzi, C. Malings, A. Minca, Information avoidance and overvaluation under epistemic constraints: Principles and implications for regulatory policies, Reliability Engineering and System Safety 197 (2020) 106814.
Annex: A FORM approximation for the reliability-based design case

We postulate the following design limit-state function with design parameter \( a \):

\[
g_d(X, a) = g(X) + a. \tag{47}
\]

We let \( \beta_0 \) denote the FORM reliability index for \( a = 0 \) and \( \alpha = [\alpha_1, \ldots, \alpha_n] \) the corresponding row vector of FORM sensitivities. The corresponding FORM approximation \( G_1(U) \) of the limit-state function in standard normal space is

\[
G_1(U) = ||\nabla G(u^*)|| (\beta_0 - \alpha U). \tag{48}
\]

\( ||\nabla G(u^*)|| \) is the norm of the gradient of limit-state function \( G \) in standard normal space, evaluated at the design point. Without loss of generality, we set \( ||\nabla G(u^*)|| = 1 \).

We assume that \( G_1(U) \) is also a reasonable approximation of \( g(X) \) when \( a \neq 0 \), which is valid for mildly non-linear problems. Hence the approximation of \( g_d(X, a) \) in standard normal space becomes

\[
G_d(U, a) = \beta_0 - \alpha U + a. \tag{49}
\]

The corresponding approximation of the probability of failure in function of \( a \) is

\[
p_F(a) \approx \Phi(-\beta_0 - a). \tag{50}
\]

Additionally conditioning on \( u_i \) results in (by analogy with Eq. 24)

\[
p_F(u_i, a) \approx \Phi\left(\frac{\alpha_i u_i - \beta_0 - a}{\sqrt{1 - \alpha_i^2}}\right). \tag{51}
\]

Using a linear cost function \( c_d(a) = c_\delta a \), the optimal design choice following Eq. 36 is

\[
a_{opt} = \arg \min_a c_\delta a + p_F(a)c_F \\
\approx \arg \min_a c_\delta a + \Phi(-\beta_0 - a) c_F. \tag{52}
\]

Without loss of generality, we choose \( c_\delta \) such that \( a_{opt} = 0 \) is the optimal design in the a-priori case, which results in

\[
c_\delta = \varphi(-\beta_0) c_F, \tag{53}
\]

wherein \( \varphi \) is the standard normal PDF.

The conditionally optimal design is

\[
a_{opt|U_i}(u_i) = \arg \min_a c_\delta a + p_F(u_i, a)c_F \\
\approx \arg \min_a c_\delta a + \Phi\left(\frac{\alpha_i u_i - \beta_0 - a}{\sqrt{1 - \alpha_i^2}}\right) c_F \\
= \alpha_i u_i - \beta_0 + \sqrt{-2 \ln \left(\frac{2\pi c_\delta}{c_F \sqrt{1 - \alpha_i^2}}\right)} \left(1 - \alpha_i^2\right). \tag{54}
\]

Inserting Eq. 53 results in

\[
a_{opt|U_i}(u_i) = \alpha_i u_i - \beta_0 + \sqrt{\beta_0^2 - \ln (1 - \alpha_i^2)} \left(1 - \alpha_i^2\right). \tag{55}
\]

For ease of notation, we introduce

\[
b = \sqrt{\beta_0^2 - \ln (1 - \alpha_i^2)}. \tag{56}
\]
By inserting the above expressions into Eq. 58, the EVPPI follows as

$$EVPPI_X = c_S a_{opt} + p_F(a_{opt}) c_F - E_{U_i} \left[c_S a_{opt|U_i}(U_i) + p_F(U_i, a_{opt|U_i}(U_i)) c_F\right]$$

$$= \Phi(-\beta_0) c_F - E_{U_i} \left[c_S \left(\alpha_i u_i - \beta_0 + b \sqrt{1 - \alpha_i^2}\right) + \Phi(-b) c_F\right]$$

$$= \left[\Phi(-\beta_0) + \varphi(-\beta_0) \left(\beta_0 - b \sqrt{1 - \alpha_i^2}\right) - \Phi(-b)\right] c_F,$$

where we have utilized $a_{opt} = 0$ and $E_{U_i}[U_i] = 0$.

Figure 13 summarizes the resulting EVPPI in function of $\alpha_i^2$ for different $\beta_0$. The relative EVPPI, which is obtained by dividing the EVPPI with the value obtained at $\alpha_i^2 = 1$, is almost identical for different $\beta_0$, as shown in Figure 8. Note that the relative EVPPI does not depend on $c_F$ and $c_S$. Hence the relative EVPPI values shown in Figure 8 are expected to be approximately valid for any reliability problem, which has a single design point and a limit-state function that is mildly non-linear in standard normal space.