Structural Safety 60 (2016) 1–15

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

# Reliability analysis with Metamodel Line Sampling

Ivan Depina [a,*], Thi Minh Hue Le [b,c], Gordon Fenton [d], Gudmund Eiksund [a]

[a] Norwegian University of Science and Technology, Department of Civil and Transport Engineering, Hogskoleringen 7A, 7491 Trondheim, Norway

[b] SAMCoT – Sustainable Arctic Marine and Coastal Technology, Centre for Research-based Innovation, Norwegian University of Science and Technology, Hogskoleringen 7A, 7491 Trondheim, Norway

[c] Klima2050 – Risk Reduction through Climate Adaptation of Building and Infrastructure, Centre for Research-based Innovation, SINTEF Building and Infrastructure, Trondheim 7034, Norway

[d] Dalhousie University, Department of Engineering Mathematics, 1360 Barrington Street, Halifax, Canada

ARTICLE INFO

*Article history:*
Received 10 December 2014
Received in revised form 22 December 2015
Accepted 27 December 2015
Available online 21 January 2016

*Keywords:*
Metamodel
Reliability
Line Sampling
Kriging
Monopile
Offshore

ABSTRACT

This paper presents an approach for reliability analysis of engineering structures, referred to as Metamodel Line Sampling (MLS). The approach utilizes a metamodel of the performance function, within the framework of the Line Sampling method, to reduce computational demands associated with the reliability analysis of engineering structures. Given a metamodel of the performance function, the failure probability is estimated as a product of a metamodel-based failure probability and a correction coefficient. The correction coefficient accounts for the error in the metamodel estimate of failure probability introduced by the replacement of the performance function with a metamodel. Computational efficiency and accuracy of the MLS approach are evaluated with the Kriging metamodel on analytical reliability problems and a practical reliability problem of a monopile foundation for offshore wind turbine. The MLS approach demonstrated efficient performance in low to medium-dimensional reliability problems.


## 1. Introduction

Reliability analysis is performed to address the inherent randomness of structural parameters and a lack of knowledge about the driving processes defining the behavior of structures. A primary interest in reliability analysis of structures is to evaluate the probability of unsafe or undesired state of the structure, i.e., failure probability, $P_F$. Given an $n$-dimensional vector of random variables affecting the performance of a structure, $\mathbf{Z} = [Z_1, \ldots, Z_n]^T \in \Omega$, in the variable space $\Omega$, associated with the joint probability density function (*pdf*), $f_{\mathbf{Z}}(\mathbf{z}), P_F$ is defined as:

$$P_F = P(\mathbf{Z} \in F) = \int_F f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} = \int_{\mathbb{R}^n} I_F(\mathbf{z}) f_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z} \tag{1}$$

where $\mathbf{z} \in \mathbb{R}^n$ denotes a realization of $\mathbf{Z}, F$ is the failure domain, $I_F$ is an indicator function such that $I_F(\mathbf{z}) = 1$ if $\mathbf{z} \in F$ and $I_F(\mathbf{z}) = 0$ otherwise. In this study, $\mathbf{Z}$ is defined as a vector of independent standard normal random variables with the joint *pdf* $\phi_{\mathbf{Z}}$, in the standard normal space $\Omega$. In the case of a general random vector $\mathbf{X}$, composed of non-normal and dependent random variables, it is assumed that a probability preserving transformation, $\mathbf{Z} = \Theta_{\mathbf{X},\mathbf{Z}}(\mathbf{X})$ (e.g., Nataf [7]) exists. It is worth noting that the transformations to the standard normal space are often approximate and can introduce additional nonlinearities in the shape of the failure domain.

The state of a structure or an engineering system is commonly evaluated by a so-called performance function, $g(\mathbf{z})$. $g(\mathbf{z})$ plays a central role in the reliability analysis of structures, because it separates the $n$-dimensional variable space $\Omega$ into a safe $g(\mathbf{z}) > 0$, and an unsafe domain $\{\mathbf{z} \in F \subset \mathbb{R}^n : g(\mathbf{z}) \leqslant 0\}$ by the hypersurface denoted as the failure limit state $\{\mathbf{z} \in L : g(\mathbf{z}) = 0\}$. In the majority of applications $g(\mathbf{z})$ is an implicit function of the random structural parameters, $\mathbf{z}$, (e.g., finite element model). The implicit formulation of the performance function introduces constraints on the applicable mathematical tools for the evaluation of $P_F$, as often only pointwise evaluations of the performance function and its gradients are obtainable.

Analytical solutions of $P_F$ are achievable only for a limited group of problems with explicit formulations of $g(\mathbf{z})$ and simple definitions of failure domains. In reliability analysis of structures, $P_F$ is often evaluated numerically by employing optimization (e.g., First and Second Order Reliability Method) or sampling methods (e.g., Monte Carlo, Importance Sampling, Subset Simulation) [22]. Among these, the Monte Carlo (MC) method is widely used due to its straightforward implementation and robust performance [22]. The MC method is based on drawing $N$ independent identically distributed (i.i.d.) samples $\mathbf{z}_i \sim \phi_{\mathbf{Z}}(\mathbf{z});\ i = 1, \ldots, N$ and evalu-

* Corresponding author.
*E-mail address:* ivan.depina@ntnu.no (I. Depina).

http://dx.doi.org/10.1016/j.strusafe.2015.12.005
0167-4730/

ating $g(\mathbf{z}_i)$ at these samples. The unbiased estimate of the failure probability, $\hat{P}_F$, is calculated as the ratio of the number of failed samples, $N_F$, over the total number of samples, $N$:

$$\hat{P}_F = \frac{1}{N}\sum_{i=1}^{N} I_F(\mathbf{z}_i) = \frac{N_F}{N} \tag{2}$$

where $N_F$ is binomial distributed random variable, which leads to the coefficient of variation of $\hat{P}_F, \mathrm{CoV}(\hat{P}_F) \approx \sqrt{(1-\hat{P}_F)/(\hat{P}_F N)}$. Investigation of the $\mathrm{CoV}(\hat{P}_F)$ reveals that the $\hat{P}_F$ is independent of the dimensionality of the problem in the MC method, and that the $\mathrm{CoV}(\hat{P}_F)$ reduces with increasing $N$. For a small $\hat{P}_F$, a relatively large $N$ is necessary to obtain a reasonably low $\mathrm{CoV}(\hat{P}_F)$. Large numbers of simulations of $g(\mathbf{z})$ are frequently infeasible to execute as the models used to evaluate $g(\mathbf{z})$ can be computationally demanding. Although the MC method is accurate, robust and independent of the dimensionality of the reliability problem, the method is considered to be inefficient when evaluating small $P_F$ and/or when computationally intensive structural models are used to evaluate the performance function.

The previously mentioned inefficiency of the MC method has led to the development of various methods suited for the estimation of small $P_F$ in probabilistic analysis of structures. The Importance Sampling (IS) method, based on the MC approach, introduces an importance *pdf* with a relatively high density over the failure domain of the variable space (e.g., [2]). By sampling the importance *pdf*, the IS method can provide $\hat{P}_F$ with reduced computational expense when compared to the MC method [22]. A series of benchmark tests conducted in [22] showed that the IS approach is applicable in low to medium-dimensional problems ($n < 100$) with efficiency and accuracy dependent on the implementation of the method. The Line Sampling (LS) method, based on the IS approach, evaluates $P_F$ by a number of conditional one-dimensional reliability problems along an important direction, which points to the failure domain nearest to the origin of $\Omega$ [22]. Benchmark tests in [22] showed high accuracy and efficient performance of the LS method in high-dimensional problems.

An alternative method for estimating $P_F$ in high-dimensional problems is the Subset Simulation (SS) method [1]. In the SS method, $P_F$ is expressed as a product of a series of conditional failure probabilities corresponding to, prior to the analysis, unknown intermediate failure limits. The conditional failure probabilities can be selected to be relatively high (e.g., $P = 0.1$), requiring consequently a small number of samples to be evaluated accurately.

Reductions in computational demands associated with the reliability analysis of engineering structures can be also achieved by replacing $g(\mathbf{z})$ with a computationally less expensive metamodel $\tilde{g}(\mathbf{z})$. Metamodels are commonly built by implementing statistical learning methods [12] (e.g., Neural Networks [20], Support Vector Machines [3,13], regression, or Kriging [8]) on a set of observations of $g(\mathbf{z})$ in the variable space. Several metamodel implementations showed high efficiency and accuracy in low to medium-dimensional problems ($n < 100$) (e.g., [8,3]).

An approach which aims at reducing computational cost commonly associated with the reliability analysis, referred to as Metamodel Line Sampling (MLS), is presented in this study. The MLS approach combines the efficiency of the LS method with a relatively low computational cost of $\tilde{g}(\mathbf{z})$ to provide reductions in computational expenses. Given $\tilde{g}(\mathbf{z}), P_F$ is evaluated as a product of a metamodel-based failure probability and a correction coefficient. The correction coefficient accounts for the uncertainty in the metamodel-based failure probability, resulting from the replacement of $g(\mathbf{z})$ with $\tilde{g}(\mathbf{z})$. The performance of the MLS approach is evaluated on analytical reliability problems and a practical reliability problem of a monopile foundation for offshore wind turbine.

## 2. Metamodel Line Sampling

### 2.1. Line Sampling

LS is a method which formulates a reliability problem as a number of conditional one-dimensional reliability problems in the standard normal space [21]. The formulation of the LS method is based on the assumption that an important direction, $\boldsymbol{\alpha}$, can be approximated. $\boldsymbol{\alpha}$ points to the region of the failure domain nearest to the origin of $\Omega$, as illustrated in Fig. 1. An MC estimate of $P_F$ is calculated based on a number of conditional one-dimensional reliability problems along $\boldsymbol{\alpha}$. The one-dimensional reliability problems are conditioned on the MC samples from the $(n-1)$-dimensional standard normal space of random variables orthogonal to $\boldsymbol{\alpha}$. Based on the set of benchmark tests [21], it is reported that the LS method has a wide range of applications in reliability analysis of structures, except for strongly nonlinear performance functions where $\boldsymbol{\alpha}$ cannot be estimated.

Given $\boldsymbol{\alpha}$, the failure domain, $F$, can be expressed as shown in [21]:

$$F = \left\{\mathbf{z} \in \mathbb{R}^n : z_{\boldsymbol{\alpha}} \in F_{\boldsymbol{\alpha}}(z_1^{\perp}, \ldots, z_{n-1}^{\perp})\right\} \tag{3}$$

where $z_{\boldsymbol{\alpha}}$ is a realization of the random variable, $Z_{\boldsymbol{\alpha}}$, which is defined along $\boldsymbol{\alpha}, \mathbf{z}^{\perp} \in \mathbb{R}^{n-1}$ is a realization of a vector of random variables orthogonal to $\boldsymbol{\alpha}$, denoted as $\mathbf{Z}^{\perp}$, while $F_{\boldsymbol{\alpha}}$ is a function representing the failure domain along $\boldsymbol{\alpha}$, defined on $\mathbb{R}^{n-1}$ [21]. Then $P_F$ can be expressed as:

$$P_F = \int_{\mathbb{R}^n} I_F(\mathbf{z})\phi_{\mathbf{Z}}(\mathbf{z})d\mathbf{z} = E_{\mathbf{Z}^{\perp}}\left[\Phi(F_{\alpha}(\mathbf{z}^{\perp}))\right] \tag{4}$$

In the case that $F_{\boldsymbol{\alpha}}(\mathbf{z}^{\perp})$ lies within the half open interval $[\beta(\mathbf{z}^{\perp}), \infty)$, the one-dimensional conditional failure probability can be evaluated as $\Phi(F_{\alpha}(\mathbf{z}^{\perp})) = \Phi(-\beta(\mathbf{z}^{\perp}))$, where $\beta(\mathbf{z}^{\perp})$ is a 'reliability index', as indicated in Fig. 1. An unbiased estimate of $P_F$ is calculated on a set of samples $\{\mathbf{z}_i^{\perp} \sim \phi_{\mathbf{Z}^{\perp}}(\mathbf{z}^{\perp}) : \ i = 1, \ldots, N\}$ as:

$$\hat{P}_F = \frac{1}{N}\sum_{i=1}^{N}\Phi(F_{\alpha}(\mathbf{z}_i^{\perp})) = \frac{1}{N}\sum_{i=1}^{N}\Phi(-\beta(\mathbf{z}_i^{\perp})) = \frac{1}{N}\sum_{i=1}^{N}P_{Fi} \tag{5}$$

where $P_{Fi} = \Phi(-\beta(\mathbf{z}_i^{\perp}))$. Variance of the estimator $\hat{P}_F$ is estimated as:

$$\mathrm{Var}(\hat{P}_F) = \frac{1}{N(N-1)}\sum_{i=1}^{N}\left(P_{Fi} - \hat{P}_F\right)^2 \tag{6}$$

Coefficient of variation of $\hat{P}_F$, defined as $\mathrm{CoV}(\hat{P}_F) \approx \sqrt{\mathrm{Var}(\hat{P}_F)}/\hat{P}_F$, is commonly used as a convergence measure of $\hat{P}_F$.

### 2.2. Metamodel-based failure probability

As discussed in Section 1, reliability analysis of structures can be a computationally intensive and time consuming task. One of the approaches to reduce the computational demands is to approximate $g(\mathbf{z})$ with a computationally less expensive metamodel, $\tilde{g}(\mathbf{z})$. A metamodel is commonly constructed by implementing statistical learning methods on a set of observations of $g(\mathbf{z})$ obtained with an information gathering process known as Design of Experiments (DoE) (e.g., Latin Hypercube Sampling). Some of the early metamodels employed first- and second-order polynomials to approximate the limit state in the proximity of the design point (i.e., the most probable point at the limit state) (e.g., [4]). More