



ELSEVIER

Contents lists available at ScienceDirect

## Structural Safety

journal homepage: [www.elsevier.com/locate/strusafe](http://www.elsevier.com/locate/strusafe)

# Reliability-based alarm thresholds for structures analysed with the finite element method

Johan Spross<sup>a,\*</sup>, Tobias Gasch<sup>b</sup>

<sup>a</sup> Division of Soil and Rock Mechanics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

<sup>b</sup> Division of Concrete Structures, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

## ARTICLE INFO

## Keywords:

Structural reliability  
Subset simulation  
Finite element method  
Monitoring  
Alarm threshold

## ABSTRACT

Civil engineering structures are commonly monitored to assess their structural behaviour, using alarm thresholds to indicate when contingency actions are needed to improve safety. However, there is a need for guidelines on how to establish thresholds that ensure sufficient safety. This paper therefore proposes a general computational algorithm for establishment of reliability-based alarm thresholds for civil engineering structures. The algorithm is based on Subset simulation with independent-component Markov chain Monte Carlo simulation and applicable with both analytical structural models and finite element models. The reliability-based alarm thresholds can straightforwardly be used in the monitoring plans that are developed in the design phase of a construction project, in particular for sequentially loaded structures such as staged construction of embankments. With the reliability-based alarm thresholds, contingency actions will only be implemented when they are needed to satisfy the target probability of failure.

## 1. Introduction

Observation of structural behaviour is standard practice in civil engineering, in particular for structures of high importance or high risk. As the cost for sensors and other equipment reduces, more and more structures are being monitored. Examples include large dams, bridges, nuclear power facilities, and geotechnical structures such as tunnels and excavations [1–7]. The purpose can be, for example, validation of design assumptions and evaluation of need for design alterations or remedial measures to ensure structural safety or satisfactory serviceability. Observations of structural behaviour can also be used to gain information about engineering properties of existing structures in assessments of their structural safety. Additional information generally implies that uncertainties are reduced and that the calculated structural reliability is improved; thereby, costly replacement or strengthening interventions may be avoided. This principle is widely applied in reliability-based design and reliability-based safety assessments of civil infrastructure; see e.g. [8–17].

As additional information is more favourable in terms of reliability improvement when uncertainties are large, observations of structural behaviour are particularly useful in geotechnical engineering, because its construction materials—soil and rock—are created by nature, which implies that their engineering properties are largely uncertain and, in addition, may exhibit a substantial inherent spatial variability.

Consequently, geotechnical design codes particularly emphasise the need for monitoring during their construction; for example, Eurocode 7 [18] requires details of the planned monitoring to be included in the Geotechnical Design Report. Moreover, the challenge of managing large uncertainties in geotechnical engineering has spurred the development of a design method in which observation during construction is a key feature: “the observational method” [18,19].

When monitoring or other types of observation of the structural behaviour are targeting structural safety, an essential concern is how to ensure that safety-enhancing contingency actions are put into operation in time. A common method is to establish an alarm, which helps the decision maker to timely interventions based on the monitoring results, but lets the decision maker attend to other tasks most of the time [20]. When the alarm threshold is violated, the decision maker is alerted to act and failure of the monitored structure can be avoided.

Despite the crucial role of alarm thresholds to ensure structural safety and satisfactory serviceability, there is little guidance available to the designing engineer on how to establish them. For example, neither Eurocode 7 nor the available application guidelines provide any detailed advice: Frank et al. [21] point out that “it is the designer’s responsibility to prepare and communicate specifications for any such monitoring”. This lack of guidance causes problems especially when applying the observational method, as the alarm threshold defines when the design must be changed. This deficiency may have

\* Corresponding author.

E-mail addresses: [johan.spross@byv.kth.se](mailto:johan.spross@byv.kth.se) (J. Spross), [tobias.gasch@byv.kth.se](mailto:tobias.gasch@byv.kth.se) (T. Gasch).

<https://doi.org/10.1016/j.strusafe.2018.09.004>

Received 14 August 2017; Received in revised form 17 September 2018; Accepted 22 September 2018

Available online 09 October 2018

0167-4730/© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

contributed to the limited use of the observational method; the need to clarify the safety aspects of its application has been discussed for decades [22–25].

In this paper, we address this lack of guidance and discuss how to establish alarm thresholds for monitored structures so that their structural safety and serviceability is continuously satisfactory. The paper builds directly on the findings of Spross and Johansson [26], who presented a reliability-based methodology that aids a decision-making engineer in choosing between the observational method and conventional design. Their methodology also showed how alarm thresholds need to be related to the acceptable probability of failure of the structure. However, Spross and Johansson [26] mainly discussed the decision-theoretical considerations regarding the application of the observational method. Therefore, we focus this paper on the more general issue of establishing alarm thresholds for civil engineering structures. While Spross and Johansson only looked at examples with analytical solutions, we here show how their methodology can be applied to a more general class of engineering problems by making use of the Finite element method and Subset simulation.

The paper is structured as follows: Chapter 2 describes the features of an alarm that ensures structural safety; Chapter 3 provides the structural reliability considerations in the establishment of alarm thresholds; Chapter 4 presents an algorithm for how reliability-based alarm thresholds can be set for structures that are analysed with the finite element method; Chapter 5 presents an illustrative example where the algorithm is applied to a concrete beam; Chapter 6 discusses the applicability of the proposed algorithm to civil engineering structures; and Chapter 7 summarises the major findings.

## 2. What is an alarm?

Alarm as a concept may be differently defined depending on the discipline. Wallin [27] identifies three different definitions (Fig. 1). In civil engineering, the stimuli-based model is normally used, as it allows a technical definition of the alarm based on the state of the monitored object. In contrast, the response-based model implies that the observer defines what constitutes as an alarm based on the incoming information, such as when an operator at a public-safety answering point decides on whether to send the rescue service or not to the caller. The message-based model refers to cases where the term “alarm” is used for the alarm notification exchanged by systems; this is common in the telecom industry. In the context of structural safety, the stimuli-based alarm model implies that structural behaviour is monitored and when some predefined threshold is violated, the alarm goes off, requesting the decision maker to act.

A crucial aspect is the establishment of the alarm threshold. The threshold should neither be too conservative, nor be too allowing; while the former may lead to costly false alarms that reduce the credibility of

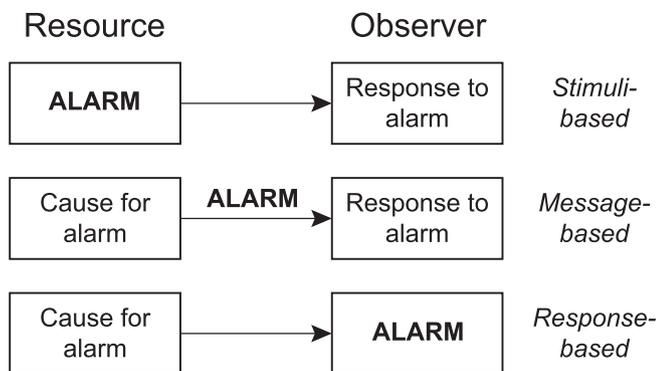


Fig. 1. Alarm definitions, extended from [59]. We use the stimuli-based definition, which is common in engineering. (© 2017. Wallin [27]. With permission of Springer.)

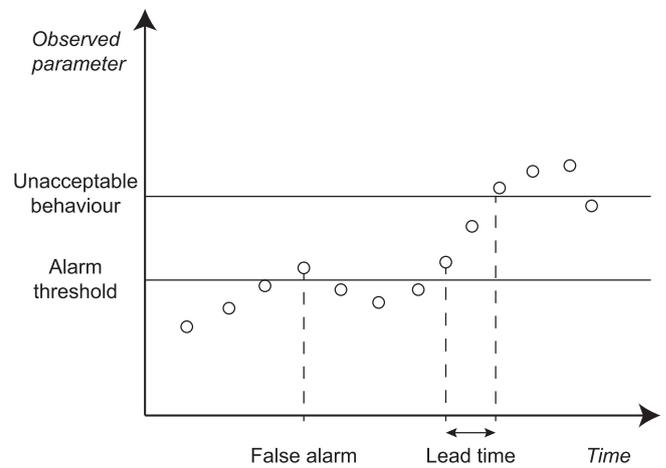


Fig. 2. False alarm and lead-time.

the alarm in the long run (known as the “cry-wolf effect”) [28–30], the latter may make the alarm go off too late, resulting in a failed structure.

The alarm threshold must be clearly distinguished from the point where unacceptable behaviour is expected to occur. The time in between the alarm threshold and the point of unacceptable behaviour is defined as the “lead-time” of the alarm (Fig. 2) [31]. This timeframe must be large enough to allow for contingency actions to be put into operation. Consequently, the required lead-time depends on the type of intervention, equipment availability, and—not to forget—the efficiency of the project organisation [32]. In a complete analysis of the lead-time, the expected failure type also needs to be considered, as the failure type will affect the available timeframe; in principle, the potential situation can be considered either time variant or time invariant. Time-variant loads either follow a more or less predictable pattern or occur as a completely unpredictable (e.g. accidental) event. For predictable load variations, the concept of lead-time is relevant; however, for completely unpredictable load increasing events, the required lead-time is by definition not possible to define. Deterioration is similar to time-variant loads, but implies instead a decrease in capacity with time. For time-invariant loads, on the other hand, any load increase is under human control and there is no restriction in time when putting contingency actions into operation. A typical example of a time-invariant load increase under human control is the decision to raise the embankment height during staged construction of road or railway embankments; additional examples are discussed in Section 6.1. Thus, in principle, the alarm threshold should be selected based on the following two aspects:

- The critical limit, where unacceptable behaviour occurs with too high probability.
- The lead-time that is required to allow for contingency actions to be put into operation.

Consequently, if a required lead-time is to be assessed accurately, the designer of the alarm system needs to consider also the possible contingency actions. This implies that all monitoring plans that involve alarm thresholds must be accompanied by a contingency action plan. The need to directly link the monitoring result to contingency actions is emphasised by Olsson and Stille [32], who suggest the following general definition of an alarm threshold in a report aiming at improving the design of the monitoring system for the construction of the Swedish nuclear waste repository:

“The alarm threshold is a predetermined value of one or a combination of several monitor parameters which, if exceeded, will trigger predetermined measures in order to prevent damage.” [Authors’ italicization]

This principle is also a key aspect of the observational method in

geotechnical engineering. In this paper, we focus on loading situations with time-invariant loads, although we discuss the possibility to extend the procedure to cover time-variant loads and deterioration in Section 6.1.

### 3. Reliability-based alarm thresholds

#### 3.1. Assessment of failure probability

In the general case, a civil engineering structure consists of one or more structural components, for which the performance can be described by the respective limit state function,  $G(\mathbf{X})$ , where the vector  $\mathbf{X}$  contains all random variables that are relevant to describe the limit state. If model errors are present, they can be straightforwardly accounted for in the limit state formulation; see e.g. [8,33]. The event of unsatisfactory performance of the component—“failure”—is defined as  $F = \{G(\mathbf{X}) \leq 0\}$ . Taking a reliability-based perspective on structural safety, unsatisfactory performance should occur with sufficiently low probability; i.e. the probability of unsatisfactory performance,  $p_F$ , should be equal to or less than the target failure probability,  $p_{F,T}$ . The relevant  $p_{F,T}$  should be clear from the applicable design guideline or code. In the general case,  $p_F$  is provided by the multidimensional integral

$$p_F = \int_{\Omega} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}, \tag{1}$$

where  $\mathbf{x}$  is the realisation of  $\mathbf{X}$ ,  $f_{\mathbf{X}}(\mathbf{x})$  is the joint probability density function of  $\mathbf{X}$ , and  $\Omega$  is the region of the failure event in the outcome space of  $\mathbf{X}$ . For component failure,

$$\Omega \equiv \{G(\mathbf{X}) \leq 0\}. \tag{2}$$

#### 3.2. Updating of failure probability with additional information

If no additional information will become available during construction, i.e. no monitoring will be performed, the safety criterion  $p_F \leq p_{F,T}$  must be satisfied with the information that is available in the design phase. This information may, for example, consist of results from site investigations, laboratory tests, experience from previous projects, and other types of engineering judgement. However, if additional information,  $Z$ , is gained at some point,  $p_F$  may be evaluated conditionally on  $Z$ . In the context of this paper, sources of information can be measurements of relevant parameters; examples include structural deformation, pore water pressure, and water inflow [10–12,25,34–36]. The updated probability of failure conditional on  $Z$  is given by

$$p_{F|Z} = \int_{\Omega_Z} f_{\mathbf{X}|Z}(\mathbf{x})d\mathbf{x}, \tag{3}$$

where the updated failure region  $\Omega_Z$  accounts for any reformulations of the limit state functions that the information called for, and  $f_{\mathbf{X}|Z}$  is the updated joint probability density function. Taking a Bayesian view on structural failure probabilities, as is common in structural reliability analysis [37], the updating can be performed with Bayes’s rule, such that

$$f_{\mathbf{X}|Z}(\mathbf{x}) = \frac{L(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})}{\int L(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}}, \tag{4}$$

where  $L(\mathbf{x})$  is the likelihood of observing  $Z$ , given the multidimensional variable  $\mathbf{X}$ .

In principle, two types of information  $Z$  from numerical measurements may be provided: either in terms of a specific measurement result (known as equality information; see [38–40]), or in terms of violated or non-violated alarm thresholds (inequality information); the latter information category is the focus of this paper.

In general terms, inequality information may be described as

$$Z = \{h(\mathbf{X}) \leq 0\}, \tag{5}$$

where the function  $h(\mathbf{X})$  defines how the structural model relates to the measurement data.

With inequality information available, the updating in Eqs. (3) and (4) becomes straightforward, because the explicit computation of  $L(\mathbf{x})$  can be circumvented: the conditional  $p_{F|Z}$  may be obtained from the definition of conditional probability,

$$p_{F|Z} = \frac{P(F \cap Z)}{P(Z)}, \tag{6}$$

because  $h(\mathbf{X})$  can be seen as a limit state function that describes the event  $Z$ . With this formulation,  $p_{F|Z}$  can be obtained with any structural reliability method, because Eq. (6) may be reformulated into

$$p_{F|Z} = P(G(\mathbf{X}) \leq 0 | h(\mathbf{X}) \leq 0) = \frac{P(\{G(\mathbf{X}) \leq 0\} \cap \{h(\mathbf{X}) \leq 0\})}{P(\{h(\mathbf{X}) \leq 0\})}, \tag{7}$$

where the numerator is analysed as a parallel-system multiple failure mode and the denominator as a single failure mode. Significant measurement error in the observations reduces the effectiveness of the updating [41]; however, this aspect is not accounted for here.

#### 3.3. Establishment of alarm thresholds based on target failure probability

Monitoring of structural behaviour and comparing the measurement result with an alarm threshold is an example of collecting inequality information. Observing that a property  $X_1$  is less than the corresponding predefined alarm threshold,  $x_{1,alarm}$ , provides the information that  $h(\mathbf{X}) = x_{1,alarm} - X_1 \leq 0$ . (For convenience, we presume that critical behaviour corresponds to exceeded thresholds; note, however, that for some applications, it may be relevant to define an alarm threshold against too low readings rather than too high. This case is straightforwardly managed by instead defining  $h(\mathbf{X}) = X_2 - x_{2,alarm} \leq 0$ , with reference to Fig. 3.)

Using Eq. (7), a value for  $x_{1,alarm}$  can be established in advance with the following equality, which ensures that the structural behaviour is acceptable (i.e.  $p_F \leq p_{F,T}$ ) as long as the monitoring result falls below the alarm threshold:

$$P(G(\mathbf{X}) \leq 0 | X_1 \leq x_{1,alarm}) = p_{F,T}. \tag{8}$$

This procedure is suggested by Spross and Johansson [26] for

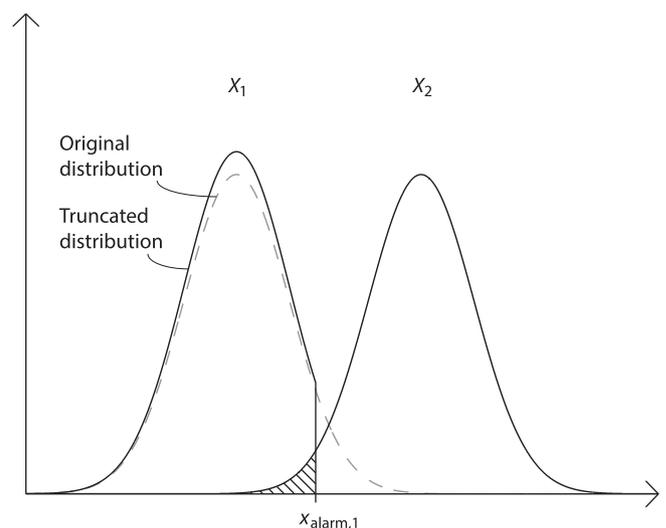


Fig. 3. An alarm threshold against too high readings of  $X_1$  is provided by a truncation of  $X_1$  at its upper end. This implies that the overlapping of the two distributions is significantly reduced and so is the corresponding failure probability. Correspondingly, an alarm threshold against too low readings of  $X_2$  may be provided if  $X_2$  is truncated at its lower end. (© 2017. Spross and Johansson [26]. CC-BY 4.0, <https://creativecommons.org/licenses/by/4.0/>.)

establishing acceptable limits of behaviour for the observational method and is illustrated in Fig. 3. An algorithm for finding  $x_{i,alarm}$  in Eq. (8) for structures using the Finite element method and Subset simulation is presented in Chapter 4. Note that in the general case, there may be more than one monitored parameter for which an alarm threshold needs to be established; i.e. we have that

$$P(G(\mathbf{X}) \leq 0 | \mathbf{X} \leq \mathbf{x}_{alarm}) = p_{F,T}, \quad (9)$$

where the vector  $\mathbf{x}_{alarm}$  contains the respective alarm thresholds for the monitored parameters. This definition implies that the alarm goes off when one of the thresholds is violated. Eq. (9) may be solved similarly to Eq. (8), but the equation becomes underdetermined when there is more than one alarm threshold to establish. Thus, to find the optimal set of alarm thresholds in  $\mathbf{x}_{alarm}$ , the cost and effectiveness of the respective measurement system need to be considered; for example, if the measurement of one parameter is expensive or associated with a large measurement error, it may be favourable not to measure this parameter at all. In essence, this is a decision-theoretical problem.

#### 4. Proposed algorithm

The advantage of using the finite element method, compared to using analytical models, to analyse civil engineering structures is that it allows the engineer both to account for more complex geometries and to capture local effects. However, the finite element analysis easily becomes computationally demanding. This poses a challenge when the finite element analysis is to be combined with a reliability analysis, because standard numerical reliability methods typically require a large number of realisations. For example, the required number of samples in a crude Monte Carlo simulation is inversely proportional to the calculated  $p_F$ . To overcome this inefficiency, more advanced Monte Carlo simulation methods have lately been developed; examples include Subset simulation [42], Line sampling [43], and Asymptotic sampling [44].

In the following, a general computational algorithm for establishing reliability-based alarm thresholds is proposed. The algorithm combines finite element analysis, to establish the limit state function, and subset simulation, to determine the failure probabilities. The theoretical foundations of these two methods are introduced, after which the algorithm is presented.

##### 4.1. The finite element method for solid mechanics

Before presenting the finite element system of equations, we establish the basic equations of solid mechanics and introduce the relevant assumptions (note the use of tensor notation in Eqs. (10) to (15)). Given the assumption of small displacements and rotations, the engineering strain definition to relate the displacement vector  $\mathbf{u}$  to the symmetric second-order strain tensor  $\boldsymbol{\varepsilon}$  is used:

$$\boldsymbol{\varepsilon} = \frac{1}{2} [(\nabla \mathbf{u})^T + \nabla \mathbf{u}] = \nabla^S \mathbf{u}, \quad (10)$$

where  $\nabla$  is the gradient operator, and  $\nabla^S$  is introduced as a notation for the symmetric gradient operator. The equilibrium equations for solid mechanics follow from Newton's second law, which, if neglecting inertial terms, state that body forces in vector  $\mathbf{f}_b$  acting over an infinitesimal volume must be balanced by the change in stress given by the second-order symmetric tensor  $\boldsymbol{\sigma}$ . This must strictly hold for every point in domain  $\Lambda$ ; i.e., we have that

$$\mathbf{0} = \nabla \cdot \boldsymbol{\sigma} + \mathbf{f}_b \text{ in } \Lambda. \quad (11)$$

Note that no difference is made between different frames of reference and, hence, neither between different stress measures. To relate the stresses and strains, Hooke's law is used:

$$\boldsymbol{\sigma} = \mathbf{C}(E, \nu): \boldsymbol{\varepsilon}, \quad (12)$$

where  $\mathbf{C}(E, \nu)$  is the fourth order elasticity tensor, which for an isotropic material is given by two material constants, e.g. Young's modulus,  $E$ , and Poisson's ratio,  $\nu$ . To complete the initial-boundary value problem, Eqs. (10) to (12) need to be complemented with appropriate initial and boundary conditions. Given that we are only interested in stationary solutions to Eq. (11), initial conditions are always given as zero displacement. Boundary conditions can be set as Dirichlet type on surface  $\Gamma_1$  by directly prescribing the displacements to be  $\mathbf{u}_0$ :

$$\mathbf{u} = \mathbf{u}_0 \text{ on } \Gamma_1. \quad (13)$$

Alternatively, Neumann type boundary conditions can be used by prescribing the surface traction  $\mathbf{t}$  on surface  $\Gamma_2$ , such that

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t} \text{ on } \Gamma_2, \quad (14)$$

where  $\mathbf{n}$  is the normal vector of the surface.

This strong form description of solid mechanics implies that equilibrium must hold locally in every point of domain  $\Lambda$ . This conflicts with the approximative nature of the finite element method. To overcome this constraint, all terms in Eq. (11) are multiplied with a test function  $\tilde{v}$  and integrated over the domain  $\Lambda$ . Then, using integration-by-parts and the divergence theorem to remove second order derivatives, we arrive at

$$\mathbf{0} = \int_{\Gamma} \tilde{v} \mathbf{t} dS - \int_{\Lambda} \nabla^S \tilde{v} : \mathbf{C}(E, \nu): \nabla^S \mathbf{u} dV + \int_{\Lambda} \tilde{v} \mathbf{f}_b dV, \quad (15)$$

which now holds for all test functions, instead of as previously for every point in  $\Lambda$ , and is on a form suitable for finite element discretization. Using the Galerkin method, we approximate both approximate  $\tilde{v}$  and  $\mathbf{u}$  as a linear combination of a set of shape functions contained in the matrix  $\mathbf{N}(\mathbb{X})$ , which describes the variation in the three-dimensional space  $\mathbb{X}$  of the solution, such that

$$\mathbf{u}(\mathbb{X}) = \mathbf{N}(\mathbb{X}) \hat{\mathbf{u}} \text{ and } \tilde{v}(\mathbb{X}) = \mathbf{N}(\mathbb{X}) \hat{v}, \quad (16)$$

whereas vectors  $\hat{\mathbf{u}}$  and  $\hat{v}$  contain discrete values of the solution in so-called nodal points. Substituting Eq. (16) into Eq. (15) and switching to matrix notation, we arrive at the discretized equilibrium equation for linear solid mechanics:

$$\mathbf{0} = -\hat{v}^T \int_{\Lambda} (\mathbf{L}\mathbf{N})^T \mathbf{C}(E, \nu) \mathbf{L}\mathbf{N} dV \hat{\mathbf{u}} + \hat{v}^T \int_{\Gamma} \mathbf{N}^T \mathbf{t} dS + \hat{v}^T \int_{\Lambda} \mathbf{N}^T \mathbf{f}_b dV, \quad (17)$$

where  $\mathbf{L}$  is the matrix equivalent of  $\nabla^S$ . Since this equation must hold for every  $\hat{v}$ , we can write the final system of equations as

$$\mathbf{K} \hat{\mathbf{u}} = \mathbf{f}, \quad (18)$$

where  $\mathbf{K}$  is the system matrix (often called stiffness matrix) defined by the integral in the first term on the right-hand side of Eq. (17),  $\hat{\mathbf{u}}$  contains the unknowns, and  $\mathbf{f}$  is the load vector given by the integral in the second and third terms. In principle, Eq. (18) gives the equations for a single finite element and to arrive at a total system of equations of an entire structure, techniques to assemble several finite elements are necessary. This is omitted here, but follows standard procedures of finite element analysis [45]. However, this global system of equations will be on an identical form, as in Eq. (18), where the unknowns  $\hat{\mathbf{u}}$  in principle are solved by multiplying both sides with  $\mathbf{K}^{-1}$ .

##### 4.2. Subset simulation

This adaptive simulation method is particularly efficient when estimating low failure probabilities for limit states with many random variables [46]. The basic idea is to calculate  $p_F$  as a product of larger conditional probabilities, which are nested intermediate events, such that  $F_0 \supset F_1 \supset \dots \supset F_M$ , where  $F_0$  is a certain event and  $F_M = F$  is the event of interest. Then, we have that

$$p_F = P\left(\bigcap_{k=1}^M F_k\right) = \prod_{k=1}^M P(F_k | F_{k-1}). \quad (19)$$

The intermediate events are defined as  $F_k = \{G(\mathbf{X}) \leq c_k\}$ , where the values of the limit  $c_k$  are set to correspond to a predefined probability  $p_0$  for the event  $F_k$ . The procedure is to simulate  $N$  samples of  $\mathbf{X}$  conditionally on the previous intermediate event,  $F_{k-1}$ , and evaluate the limit state function  $G(\mathbf{X})$  with the simulated samples. The limit  $c_k$  is set as the  $p_0$ -percentile of the calculated values of  $G(\mathbf{X})$ . To generate the  $N$  samples, independent-component Markov chain Monte Carlo simulation (MCMC) [47] is applied, using the samples that satisfied the previous intermediate event,  $F_{k-1}$ , as seeds in the simulation. “Independent-component” implies that a Metropolis–Hastings algorithm is used to generate samples from a proposal probability density function  $q_i^*$  for each parameter  $i$  in  $\mathbf{X}$  independently of each other, after which the conditioning event is checked. The number of seeds (i.e. the number of Markov chains) is  $N_c = p_0 N$  and each seed generates  $N_s = 1/p_0$  sample proposals, which then are either accepted or rejected depending on their suitability for the target distribution. The procedure is repeated until  $c_k$  becomes negative, at which point the final subset level  $M$  is reached. For the last event,  $F_M = \{G(\mathbf{X}) \leq c_M\}$ , the limit  $c_M = 0$ . The  $p_F$  can then be estimated as

$$p_F \approx \hat{p}_F = p_0^{M-1} p_M, \quad (20)$$

where  $p_M$  is the estimate of the last conditional probability  $P(F_M | F_{M-1})$  and given by

$$p_M = \frac{1}{N} \sum_{i=1}^N I_F(\mathbf{x}_{M-1}), \quad (21)$$

where  $I_F$  is the indicator function of  $F$  evaluated with the samples in  $\mathbf{x}_{M-1}$  generated conditionally on the event  $F_{M-1}$ ; i.e., we have that  $I_F(\mathbf{x}_{M-1}) = 1$ , if  $G(\mathbf{x}_{M-1}) \leq 0$  and  $I_F(\mathbf{x}_{M-1}) = 0$  otherwise. More detailed discussions of Subset simulation algorithms and their application to civil engineering problems are found in [47–49].

To establish the alarm threshold in Eq. (8),  $p_F$  must, however, be estimated conditionally on the information  $Z = \{X_1 \leq x_{1,\text{alarm}}\}$ . Following the approach suggested by Straub et al. [46], the intermediate events are reformulated into  $F_0^* \supset F_1^* \supset \dots \supset F_M^*$ , where  $F_k^* = F_k \cap Z$  and  $F_0^* = F_0 \cap Z = Z$ . Then, we can modify Eq. (19) into

$$p_{F|Z} = \frac{P(F \cap Z)}{P(Z)} = P\left(\bigcap_{k=1}^M F_k^* \mid F_0^*\right) = \prod_{k=1}^M P(F_k^* | F_{k-1}^*). \quad (22)$$

### 4.3. Simulation algorithm for establishment of alarm thresholds

The proposed algorithm for establishment of reliability-based alarm thresholds for civil engineering structures combines the concepts of Subset simulation with independent-component MCMC, structural reliability analysis, and finite element analysis. It is presumed that one of the random parameters,  $X_1$ , relevant for the limit state  $G(\mathbf{X})$  is monitored.

In principle, the algorithm uses an iterative process to find the  $x_{1,\text{alarm}}$  that satisfies the equality of Eq. (8), i.e. the  $x_{1,\text{alarm}}$  that is required to ensure that  $p_{F|Z} \leq p_{F,T}$ , which shows that the structural behaviour is acceptable. In each iteration, a Subset simulation with independent-component MCMC [47] is performed.

The iterations end when the error,  $\hat{\varepsilon}$ , between  $p_{F|Z}$  and  $p_{F,T}$  is less than a predefined tolerance,  $\tau$ , for the  $x_{1,\text{alarm}}^{(j)}$  that was suggested in the previous iteration. The algorithm is developed to allow evaluation of one or more of the needed random variables in  $\mathbf{X}$  with a finite element model, using available knowledge in terms of “basic” random variables in a vector  $\mathbf{B}$ . This can be exemplified with the need for evaluating the response of a structural component from a loading with a finite element model. In principle, though, the algorithm will work also for other types of models; hence, it is referred to as a structural model in the algorithm. In the calculation example in Chapter 5, the algorithm is applied in combination with both a finite element model and an analytical model to allow for comparison. The general algorithm has 4 main steps and is

**Table 1**

Proposed algorithm for establishment of reliability-based alarm thresholds.

1. Definition of simulation constants and basic, case-specific data.
  - a. Select  $p_{F,T}$ ,  $N$ ,  $p_0$ ,  $\tau$ , and  $\kappa$ . ( $\kappa$  is a factor to provide sufficient amount of samples after truncation in step 3.b.)
  - b. Define  $G(\mathbf{X})$  and the joint probability density function of the basic random variables,  $f_{\mathbf{B}}$ .
2. Initial crude Monte Carlo simulation.
  - a. Generate  $N\kappa$  i.i.d. sets of samples,  $\mathbf{b}$ , from  $f_{\mathbf{B}}$ .
  - b. Run the structural model (i.e., in principle, solve Eq. (18)) for each parameter set of samples in  $\mathbf{b}$  to evaluate the remaining parameters in  $\mathbf{x}$ .
  - c. Evaluate  $G(\mathbf{x})$ .
  - d. Set the first guess of the alarm threshold,  $x_{1,\text{alarm}}^{(1)}$  (see section 4.4.2 for details).
  - e.  $j = 1$ .
3. While  $\hat{\varepsilon} > \tau$  (iterative loop to find  $x_{1,\text{alarm}}$ )
  - a.  $k = 1$ .
  - b. Satisfy the initial event  $F_0^* = Z$  by accepting, from the  $N\kappa$  parameter sets in  $\mathbf{x}$ ,  $N$  sets of samples that satisfies  $x_1 < x_{1,\text{alarm}}^{(j)}$ , into a matrix  $\mathbf{x}_Z$ . If the number of acceptable sample sets is less than  $N$ , increase  $\kappa$  and start over from step 2.a.
  - c. Order the sets of samples in  $\mathbf{x}_Z$  in increasing order of magnitude of their limit state value  $G(\mathbf{x})$  and let  $c_k$  be the  $p_0$ -percentile of the ordered samples. Let the  $N_c = p_0 N$  first parameter sets of samples be denoted  $\mathbf{x}_k^{(\text{seed})}$ , while  $\mathbf{b}_k^{(\text{seed})}$  contains the corresponding seeds of the Markov chains for the basic random variables. Set  $F_k^* = \{G(\mathbf{x}_Z) \leq c_k\}$ .
  - d. While  $c_k > 0$  (Iterative loop for Subset simulation)
    - i. For all  $N_c$  Markov chains (NB: index  $k$  is suppressed in step i. for convenience):
      - Generate  $N_s = 1/p_0$  sets of conditional samples  $\tilde{\mathbf{b}}_l = [\tilde{b}_1^{(l)}, \tilde{b}_2^{(l)}, \dots, \tilde{b}_n^{(l)}]$  from a proposal PDF  $q_i^*(b_i^{(\text{seed})})$  for each basic parameter in  $\mathbf{B}$ .
      - Calculate for each proposal  $r_l^{(i)} = \frac{q_i^*(\tilde{b}_1^{(l)}, \tilde{b}_2^{(l)}, \dots, \tilde{b}_n^{(l)}) f_{\mathbf{B}}(\tilde{\mathbf{b}}_l^{(i)})}{q_i^*(b_1^{(\text{seed})}, b_2^{(\text{seed})}, \dots, b_n^{(\text{seed})})}$ .
      - Generate  $u_l^{(i)}$  uniformly distributed on  $[0, 1]$  for each proposal.
      - Set for all  $l$  and all basic parameters,  $\tilde{b}_l^{(i)} = \begin{cases} \tilde{b}_l^{(i)} & \text{if } u_l^{(i)} < r_l^{(i)} \\ b_l^{(\text{seed})} & \text{otherwise} \end{cases}$ .
    - ii. Run the structural model for each set of samples in  $\tilde{\mathbf{b}}_k$  to find the complete set of proposed samples,  $\tilde{\mathbf{x}}_k$ .
    - iii. Set, for all  $N$  parameter sets of proposed samples,  $\mathbf{x}_k = \begin{cases} \tilde{\mathbf{x}}_k & \text{if } \tilde{\mathbf{x}}_k \in F_k^* \\ \mathbf{x}_k^{(\text{seed})} & \text{otherwise} \end{cases}$ .
    - iv. Order the samples in  $\mathbf{x}_k$  in increasing order of magnitude of their limit state value  $G(\mathbf{x})$  and let  $c_{k+1}$  be the  $p_0$ -percentile of the ordered samples. Let the  $N_c$  first samples be denoted  $\mathbf{x}_{k+1}^{(\text{seed})}$  and let the corresponding  $\mathbf{b}_{k+1}^{(\text{seed})}$  contain the next seeds of the Markov chains.
    - v.  $k = k + 1$ .
  - e. Identify the number,  $N_F$ , of sample sets for which  $\mathbf{x}_{k-1} \in F_M^*$  and calculate  $\hat{p}_F = p_0^{k-1} \frac{N_F}{N}$ .
  - f. Calculate the error between  $\hat{p}_F$  and  $p_{F,T}$  as  $\hat{\varepsilon} = \frac{|\hat{p}_F - p_{F,T}|}{p_{F,T}}$ .
  - g. Set  $x_{1,\text{alarm}}^{(j+1)}$  to be  $\begin{cases} < x_{1,\text{alarm}}^{(j)} & \text{if } \hat{p}_F > p_{F,T}(1 + \tau) \\ > x_{1,\text{alarm}}^{(j)} & \text{if } \hat{p}_F < p_{F,T}(1 - \tau) \end{cases}$  (see Section 4.4.2 for details).
  - h.  $j = j + 1$ .
  4.  $x_{1,\text{alarm}} = x_{1,\text{alarm}}^{(j-1)}$ .

presented in Table 1.

### 4.4. Comments on the algorithm

#### 4.4.1. Definition of simulation constants

The constant  $p_0$  determines the intermediate probabilities and affects how many seeds that are picked out in step 3.c. According to Zuev et al. [50],  $p_0$  should be set in the range  $[0.1, 0.3]$  to ensure high efficiency.  $N$  should be selected large enough to estimate  $p_0$  accurately [48]. Also, it must be ensured that  $N_c = p_0 N$  and  $N_s = 1/p_0$  are positive integers. The constant  $\kappa$  ensures that there are enough samples entering the truncation procedure in step 3.b. to have  $N$  parameter sets of samples in  $\mathbf{x}_Z$ . Consequently, the required  $\kappa$  will depend on the number of rejected sets of samples in the truncation. The choice of  $\tau$  adjusts the accuracy in the final  $p_{F,T}$  for the accepted  $x_{1,\text{alarm}}$  in step 3.g.

4.4.2. Iterative suggestions of the alarm threshold

In step 2.d., the first suggestion of the alarm threshold,  $x_{1,alarm}^{(1)}$ , is made. In principle,  $x_{1,alarm}^{(1)}$  can be set anywhere in the generated sample  $x_1$ . For load-related parameters, we suggest to set  $x_{1,alarm}^{(1)}$  in the range  $[\bar{x}_1, \max x_1]$ , where  $\bar{x}_1$  is the sample mean, and for resistance-related parameters, we suggest the range  $[\min x_1, \bar{x}_1]$ . Note that if the alarm threshold is needed in the other end of the range of  $x_1$  than suggested above, in order to satisfy  $p_{F,T}$ , the probability of violating  $x_{1,alarm}$ , i.e.  $P(X_1 > x_{1,alarm})$  for load-related parameters, will be substantial. This would indicate a high probability of needing to put contingency actions into operation to ensure structural safety, as discussed by Spross and Johansson [26].

In step 3.g., a new guess is to be made for  $x_{1,alarm}^{(j)}$ . As the calculated  $\hat{p}_F$  indicates whether  $x_{1,alarm}^{(j)}$  is set too high or too low, each iteration will reduce the possible range of the final  $x_{1,alarm}$ . Therefore, we suggest that each new guess should be aimed to reduce the remaining range as much as possible. A simple solution is to set  $x_{1,alarm}^{(j+1)}$  to be the average value of maximum and minimum values of the remaining possible range.

4.4.3. Proposal PDF in the MCMC

The selection of proposal PDFs  $q_i^*$  that are used to generate the candidate samples  $\tilde{b}_i$  in step 3.d.i. will affect the transition from the current state to the next. As discussed by Au and Wang [47], a safe strategy, as well as a convenient choice, is to make a random shift from the current sample (a “Metropolis random walk”), letting  $\tilde{b}_i$  be normally distributed with mean  $\tilde{\mu}_{b,i} = b_i^{(seed)}$  and standard deviation  $\tilde{\sigma}_{b,i} = \sigma_{b,i}$ , which is the standard deviation of  $f_{b_i}$ . Other possibilities include using a uniform or triangular distribution.

4.4.4. Efficiency of subset simulation algorithms

Our algorithm is based on the algorithm for subset simulation with independent-component MCMC proposed in [47], which we find sufficiently effective for the problem at hand; the most time-consuming part of the algorithm lies in the evaluation of the finite element model and not in the subset simulation. However, for increased efficiency of the subset simulation, the user may straightforwardly implement recent findings on subset simulation, e.g. [48,51,52], in their application.

5. Illustrative example

5.1. Case description: Deformation of a simply supported beam

To illustrate the proposed algorithm for establishment of alarm thresholds, we have applied it to find a reliability-based alarm threshold for a simply supported reinforced concrete beam loaded with a distributed load (Fig. 4). The threshold indicates when the probability of violating the limit state becomes unacceptably high. This simplified case allows evaluation of the computational cost of using a finite

**Table 2**  
Random properties for the analysed concrete beam.

Parameter	Symbol	Distribution type	Mean	Coefficient of variation
Tensile strength	$f_t$	Lognormal	3 MPa	0.15
Young’s modulus	$E$	Lognormal	30 GPa	0.15
Distributed area load	$q$	Lognormal	6 kN/m <sup>2</sup>	0.15

element model, as there is an analytical solution available. Concrete is chosen as the material for the beam; however, the algorithm will work for any material. Random properties for the basic random variables in  $B$  are chosen as the tensile strength  $f_t$ , Young’s modulus  $E$  and the distributed area load  $q$ ; their distributions are presented in Table 2. The dimensions of the beam are fixed (Fig. 4) and Poisson’s ratio  $\nu$  is set to 0.2. For simplicity, we have not considered spatial variation of random parameters in the analysis; though, such aspects may straightforwardly be added to the finite element model.

The limit state function for the example is

$$G(X) = \frac{f_t}{E} - \varepsilon \leq 0, \tag{23}$$

where  $\varepsilon$  is the maximum strain in the beam subjected to  $q$ . For a simply supported beam,  $\varepsilon$  is always located at the bottom material fibre in the mid-section. The sought alarm threshold in terms of strain is denoted  $\varepsilon_{alarm}$ , which straightforwardly may be converted to an alarm threshold in terms of vertical deformation,  $u_{alarm}$ . Note that the defined limit state does not describe failure of a concrete beam; rather, it indicates when cracking first occurs, which can be seen as a serviceability limit state. This implies that the finite element analysis can be limited to linear solutions (as described in Section 4.1); however, the proposed algorithm is also applicable with nonlinear structural models, but at a significantly larger computational cost.

5.2. Analytical model

To set up the analytical solution for  $\varepsilon$ , we define from structural mechanics the maximum field moment for a simply supported beam as

$$M_f = \frac{ql^2}{8d}, \tag{24}$$

where  $l$  is the length and  $d$  the depth of the beam. With Eq. (24), the bending strain is calculated to

$$\varepsilon = \frac{M_f h}{EI2} = \frac{ql^2 h}{16EI d}, \tag{25}$$

where  $I$  is the second moment of area around the bending axis, and  $h$  is the height of the beam. The vertical deflection at mid-span is

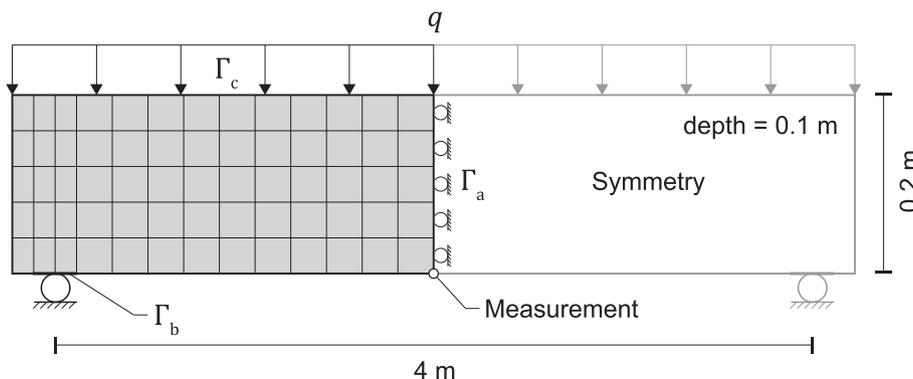


Fig. 4. Schematic description of the analysed concrete beam including dimensions, finite element discretization, and boundary conditions.

$$u = \frac{5ql^4}{384EI d} \tag{26}$$

### 5.3. Finite element model

A continuum description of the beam is set up in two-dimensions using the plane stress assumption. As indicated in Fig. 4, a vertical symmetry plane is utilized in the mid-section to reduce the model size and computational time. This is done using the Dirichlet condition in Eq. (13) on surface  $\Gamma_a$ , where the displacement component normal to  $\Gamma_a$  is constrained to zero. The roller boundary condition of surface  $\Gamma_b$  is described using a special form of the Dirichlet condition to allow for rotation of the surface around its midpoint  $\mathbb{X}_m$ . This can be written as

$$\mathbf{u} = \mathbf{u}_0 + (\mathbf{R} - \mathbf{I}) \cdot (\mathbb{X} - \mathbb{X}_m) \text{ on } \Gamma_b, \tag{27}$$

where  $\mathbf{I}$  is the unit matrix and matrix  $\mathbf{R}$  describes the rigid body rotation of  $\Gamma_b$ . The vertical component of rigid body displacement  $\mathbf{u}_0$  is constrained to zero. The distributed load is applied using the Neumann condition in Eq. (14) on surface  $\Gamma_c$ , where the normal surface traction  $t_n = -q$ . The finite element discretization of the beam is shown schematically in Fig. 4, with the element size set to  $0.04 \times 0.04$  m. Each finite element is in the example defined using quadratic serendipity shape functions and thus  $\mathbf{N}(\mathbb{X})$  becomes a  $16 \times 16$  matrix. In total, the model contains 974 unknowns. In the current study, the general-purpose finite element software COMSOL Multiphysics 5.3 [53] is used to solve the finite element model.

### 5.4. Simulation constants

Table 3 shows the values of the simulation constants that were used in the simulations. Setting  $p_{F,T} = 0.001$  corresponds to a serviceability limit state associated with high consequence, as proposed by Fenton et al. [54]. For the Subset simulation, 10 000 sets of samples were generated in step 2.a (see Table 1), out of which the first 5000 that satisfied the initial event  $F_0^* = Z$  (step 3.b) were used in the subsequent simulation. To allow for comparison of the computational cost when using the respective model to evaluate the structural response, each iteration of the loop to find  $\varepsilon_{\text{alarm}}$  (step 3) was clocked. The Subset simulation algorithm was also compared to crude Monte Carlo simulation, for which  $N_{MC} = 150\,000$  sets of samples were generated. Out of these, approximately 2/3 satisfied the condition  $Z$  and were subsequently used to calculate the alarm threshold in a loop with the same acceptable tolerance  $\tau$  as in the Subset simulation. Matlab R2013b [55] was used for the simulations.

### 5.5. Calculation results

Fig. 5a illustrates the sample after step 3.b in the algorithm in the last iteration (Table 1). The contour lines correspond to  $F_0$  (i.e. the initial crude Monte Carlo simulation) and the black dots to  $F_0^*$ . Fig. 5b illustrates the last generated sample of the MCMC procedure after rejecting the unsuitable proposals. The calculated alarm thresholds for the respective structural model are presented in Table 4, together with the total calculation time and the average calculation time,  $\bar{t}$ , per iteration in the loop of step 3. For all cases, the loop of step 3 was

entered 4 times. As the analysed case is linearly elastic, the two structural models are expected to provide approximately the same answer. For reference, the expected probability of exceeding  $u_{\text{alarm}}$  is 28% (equivalent to the portion of grey dots in Fig. 5a) and the  $p_F$  without the alarm is 0.0043.

Timing the analyses for the two structural models and the two simulation methods illustrates the efficiency of the algorithm for the two structural models (Table 4). Using Subset simulation together with the analytical model is 29 times more effective than with the finite element model in this illustrative example. Clearly, the evaluation of the finite element model effectively stands for almost all computational effort when such structural models are used with the algorithm. Note also the long  $\bar{t}$  when using Subset simulation and a finite element model (375 s/iteration) compared to using crude Monte Carlo simulation (110 s/iteration). The reason is that Subset simulation requires  $N$  evaluations of the structural model in each iteration of the loop of step 3.d, while for crude Monte Carlo simulation the structural model is run  $N_{MC}$  times, but outside of the loops (at the equivalence of step 2.b in the proposed algorithm in Table 1). For the given set of model parameters in Table 3, the total number of limit state function calls is approximately 50 000 for the subset simulation (depending on the convergence of step 3 and step 3.d, respectively), compared to 150 000 for the crude Monte Carlo simulation.

Investigating the variability in the calculated results, we found that the calculated  $u_{\text{alarm}}$  is affected mainly by the number of Subset simulations,  $N$ . Increasing  $N$  reduces the variability in  $u_{\text{alarm}}$ ; for example, increasing  $N$  from 5000 to 10 000 implies a reduction in coefficient of variation for  $u_{\text{alarm}}$  from 4.7% to 3.1% (based on 50 calculations of  $u_{\text{alarm}}$  each). The variability reduction obviously comes with a larger computational cost.

## 6. Discussion

### 6.1. Practical applicability to civil engineering structures

As discussed in Chapter 2, the purpose of establishing an alarm threshold for a structure is to ensure that action is taken to prevent structural failure or unsatisfactory performance, e.g. in terms of violation of a limit state. The proposed algorithm has the advantage of letting the decision maker establish the threshold based on  $p_{F,T}$ , thereby ensuring that the structure is sufficiently safe as long as the threshold is not violated. Though, this implies that if a threshold is violated, so is also the  $p_{F,T}$ , which should not be acceptable. However, this is not true for all cases; it is therefore important to distinguish between the following two situations: (1) when observations are used to predict a future behaviour for which an alarm has been established, and (2) when observations are compared directly against the alarm limit to assess the current situation.

In many civil engineering projects, the former is the case. Typical examples are civil engineering structures that have a sequential load increase during construction, such as the excavation of a rock tunnel where each blast round implies additional loading on the lining [17] or the staged construction of an embankment [11]. Then, measurements from each stage can be used to predict the final behaviour, implying that the predicted final behaviour is compared against the alarm threshold rather than the current behaviour. Thus, if  $p_{F,T}$  is expected to be violated because of the next load increase, the prepared contingency actions need to be put into operation before the load increase is made. This is possible when the loading is completely under human control (i.e., time invariant).

The latter case, where measurements only relate to the current situation, can be exemplified with piezometric measurements of uplift pressure during the remedial grouting of a dam foundation; too large uplift pressure may cause sliding failure of the dam [25]. As a future piezometric pressure would be difficult to predict, the alarm threshold for when to drill relief wells to reduce the pressure must relate to the

**Table 3**  
Simulation constants used in the example.

Constant	Value
$p_{F,T}$	0.001
$p_0$	0.1
$\tau$	0.1
$N \cdot \kappa$ (Subset sim.)	5000 · 2
$N_{MC}$ (crude MC)	150 000

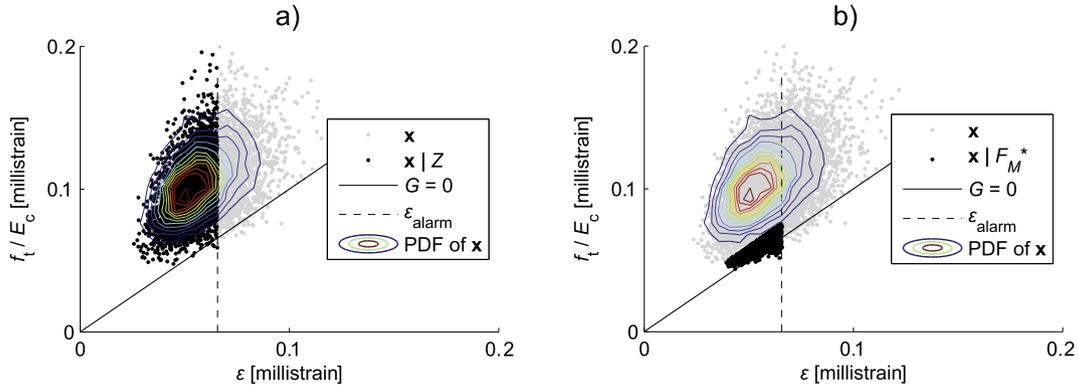


Fig. 5. Simulation results for the concrete beam. a) The remaining sample (in black), given that  $\epsilon < \epsilon_{\text{alarm}}$ . b) The last generated subset sample, given that  $G(\mathbf{X}) < c_k$  and  $\epsilon < \epsilon_{\text{alarm}}$ .

Table 4

Calculation results: alarm thresholds in terms of both bending strain and vertical deformation, as well as the total calculation time and the average calculation time for one iteration of step 3 in the algorithm (Table 1).

Parameter	Subset sim.		Crude Monte Carlo sim.	
	Analytical model	FEM	Analytical model	FEM
$\epsilon_{\text{alarm}}$ [millistrain]	0.0670	0.0678	0.0661	0.0659
$u_{\text{alarm}}$ [mm]	1.11	1.11	1.10	1.10
$t_{\text{tot}}$ [s]	61	1781	452	4543
$\bar{t}$ [s/iter. of step 3]	15	375	113	110

current situation. The threshold must therefore also account for the lead-time of this contingency action (cf. Fig. 2), such that the uplift pressure increase can be stopped before the required  $p_{F,T}$  is violated.

Application of the proposed algorithm is most straightforward to structures of the former category, i.e. when observations are used to predict a future behaviour and the alarm threshold does not need to account for a lead-time because of a time-variant load. Regarding the time-variant situations (see Section 2), the proposed procedure can be extended to cases with predictable load changes and deterioration, given that the variation with time,  $t$ , of the variables can be described in the limit state function:  $G(\mathbf{X}, t) = 0$ . Presuming full correlation between measurements of the variable  $X_1$  at the time  $t_i$  and those at the later time  $t_i + \Delta t_{\text{lead}}$  (i.e. after the potential lead-time  $\Delta t_{\text{lead}}$ , if contingency actions were put into operation at  $t_i$ ), Eq. (8) can be extended to

$$P(G(\mathbf{X}, t_i + \Delta t_{\text{lead}}) \leq 0 | X_1(t_i + \Delta t_{\text{lead}}) \leq x_{1,\text{alarm}}(t_i + \Delta t_{\text{lead}})) = p_{F,T}. \quad (28)$$

Having established  $x_{1,\text{alarm}}(t_i + \Delta t_{\text{lead}})$  with the proposed algorithm (Table 1) as a percentile  $P_{\text{alarm}}$  of  $X_1$  (cfr. Fig. 3), the sought alarm threshold for  $t_i$  is then obtained from the corresponding percentile of  $X_1$  at  $t_i$ :

$$x_{1,\text{alarm}}(t_i) = P_{\text{alarm}}(X_1(t_i)). \quad (29)$$

Conceptually, Eqs. (28) and (29) imply that the alarm goes off at  $t_i$  if the measurement of  $X_1$  is found in the same percentile as the percentile that was cut off by the alarm threshold truncation occurring when the algorithm was run for the predicted situation at  $t_i + \Delta t_{\text{lead}}$ . To establish such time-variant alarm thresholds, the structural model needs to be time dependent and have sufficient accuracy and temporal resolution to allow evaluation and interpolation for every measurement that is taken. Consequently, the procedure can quickly become very computationally expensive.

As discussed in Chapter 2, false alarms can both be costly and reduce the credibility of the alarm. In the context of structural monitoring, false alarms can be adhered to a scenario where the alarm malfunctions and goes off even though the observed structural behaviour is acceptable, e.g. because of some technical or human error. Such

situations are, in our opinion, best addressed by other means a risk management, such as quality control of the installation of the alarm system. Note that a correctly measured violation of the alarm threshold, that is not causing failure, should not be interpreted as a false alarm in the context of time-invariant loads, as the remaining margin to failure does not satisfy the required  $p_{F,T}$ .

### 6.2. Inequality or equality information?

A possible objection to using alarm thresholds to ensure structural reliability is that all information in the performed measurement is not used; i.e., if the measurement data were treated as equality information such that  $Z = \{h(\mathbf{X}) = 0\}$ , instead of treating the data as inequality information (Eq. (5)), the updated failure probability  $p_{F|Z}$  would potentially be more reduced. We believe, however, that the use of alarm thresholds has a practical advantage over direct computation of  $p_{F|Z}$ , as this allows simple offline readings on site. This makes the principle of an alarm threshold easy to communicate to the staff at the construction site, which is beneficial from a risk management perspective. Thus, reliability-based alarm thresholds can straightforwardly be implemented by the designing engineer as a part of a monitoring plan, such as the plan required by Eurocode 7 for geotechnical structures. This would be an improvement compared to today's practice, where alarm thresholds normally are established on a less rigorous analysis. The use of equality information requires, in comparison, rather sophisticated online computer computations during the course of construction before it can be decided whether a contingency action is needed or not.

### 6.3. The concept of acceptable failure probability

As the proposed algorithm uses a limit state function to establish the alarm threshold, it requires a clear definition of failure, i.e., unsatisfactory behaviour. For many civil engineering structures, this is achieved by defining a bearing capacity. However, the establishment of an alarm requires a measurable parameter; preferably, it should also be easy to observe. For many structures, deformation is such a parameter. Consequently, the limit state needs to be possible to formulate in terms of deformation (or strain); fortunately, this can often be achieved with a finite element analysis as suggested in the proposed algorithm, as this is a displacement-driven method. However, in rock engineering, for example, some types of unsatisfactory behaviour may be difficult to describe accurately with a limit state function; see e.g. [56–58].

The applicability of linear analyses in establishing alarm thresholds for civil engineering structures in the ultimate limit state needs to be further studied, in particular for those made of brittle materials such as concrete and rock. For example, the linear limit state function in Eq. (23) was in the example only seen as a serviceability limit state and

would not be applicable if one were interested in the failure of the beam; a reinforced concrete beam only behaves linearly until the first crack appears and the bearing capacity can exceed this limit by orders of magnitude. To extend the structural model to cover any nonlinear behaviour, e.g. plasticity or creep, to capture a realistic failure in the ultimate limit state would require more effort in the modelling and description of the related uncertainties. In addition, this would obviously increase the computational cost in solving the structural model significantly. For brittle failures that are expected to progress quickly, it may be favourable to define unacceptable behaviour at the point where the failure progression is expected to be initiated, rather than where it is expected to end with structural collapse; as a consequence, the alarm threshold then needs to be established with a safety margin to the point where failure progression is initiated.

## 7. Conclusions

We have presented a general computational algorithm for establishment of reliability-based alarm thresholds for civil engineering structures. The algorithm is based on subset simulation with independent-component MCMC and can be used both with analytical models and finite element models to evaluate the limit state function. The alarm threshold is established such that the target failure probability is satisfied as long as the observations do not violate the threshold. The concept is mainly applicable to sequentially loaded structures, where the observations can be used to predict the final behaviour. We believe that the proposed algorithm may prove useful in preparing monitoring plans for construction projects, in particular in geotechnical engineering where observations of structural behaviour often are required during construction. Contingency actions are then only implemented when they are needed to satisfy the target probability of failure.

## Acknowledgements

The presented research was funded and supported by the Rock Engineering Research Foundation (BeFo; grant no. 381). The research was conducted without involvement of the funding source.

## References

- Brownjohn JMW. Structural health monitoring of civil infrastructure. *Philos Trans R Soc A* 2007;365:589–622.
- Paté-Cornell ME, Tagaras G. Risk costs for new dams: economic analysis and effects of monitoring. *Water Resour Res* 1986;22:5–14.
- De Sortis A, Paoliani P. Statistical analysis and structural identification in concrete dam monitoring. *Eng Struct* 2007;29:110–20.
- Glišić B, Inaudi D, Vurpillot S. *Structural monitoring of concrete structures*. Como, Italy: Wiley; 2002. p. 1–10.
- Buckby RJ, Key CL, Palmer MJ, Penman JG, Swannell NG. An observational approach to design of a new grout curtain at Wembley Dam. In: Winter MG, Smith DM, Eldred PJJ, Toll DG, editors. *Proceedings of the XVI ECSMGE: geotechnical engineering for infrastructure and development*. London: ICE Publishing; 2015. p. 559–654.
- Mahini SS, Moore JC, Glencross-Grant R. Monitoring timber beam bridge structural reliability in regional Australia. *J Civ Struct Health Mon* 2016;6:751–61.
- Smith LM. In-service monitoring of nuclear-safety-related structures. *Struct Engineer* 1996;74:210–1.
- Goulet J-A, Der Kiureghian A, Li B. Pre-posterior optimization of sequence of measurement and intervention actions under structural reliability constraint. *Struct Saf* 2015;52:1–9.
- Krounis A, Johansson F, Spross J, Larsson S. Influence of cohesive strength in probabilistic sliding stability reassessment of concrete dams. *J Geotech Geoenviron Eng* 2016;04016094.
- Spross J, Johansson F, Larsson S. On the use of pore pressure measurements in safety reassessments of concrete dams founded on rock. *Georisk* 2014;8:117–28.
- Müller R, Larsson S, Spross J. Multivariate stability assessment during staged construction. *Can Geotech J* 2016;53:603–18.
- Schweckendiek T, Vrouwenvelder ACWM, Calle EOF. Updating piping reliability with field performance observations. *Struct Saf* 2014;47:13–23.
- Schweckendiek T, van der Krogt MG, Teixeira A, Kanning W, Brinkman R, Rippi K. Reliability updating with survival information for dike slope stability using fragility curves. *Geotech Sp* 2017;283:494–503.
- Papaioannou I, Straub D. Learning soil parameters and updating geotechnical reliability estimates under spatial variability—theory and application to shallow foundations. *Georisk* 2017;11:116–28.
- Li H, Li S, Ou J, Li H. Reliability assessment of cable-stayed bridges based on structural health monitoring techniques. *Struct Infrastruct Eng* 2012;8:829–45.
- Luque J, Straub D. Reliability analysis and updating of deteriorating systems with dynamic Bayesian networks. *Struct Saf* 2016;62:34–46.
- Bjureland W, Spross J, Johansson F, Prästings A, Larsson S. Reliability aspects of rock tunnel design with the observational method. *Int J Rock Mech Min Sci* 2017;98:102–10.
- CEN. EN 1997-1:2004. Eurocode 7: geotechnical design – Part 1: general rules. Brussels: European Committee for Standardisation; 2004.
- Peck RB. Advantages and limitations of the observational method in applied soil mechanics. *Geotechnique* 1969;19:171–87.
- Huang C, Schachter R. Alarms for monitoring: a decision-theoretic framework. Technical Report SMI-97-0664. Stanford: Section on Medical Informatics, Stanford University School of Medicine; 1997.
- Frank R, Bauduin C, Driscoll R, Kavvas M, Krebs Ovesen N, Orr T, et al. *Designers' guide to EN 1997-1 Eurocode 7: geotechnical design – general rules*. London: Thomas Telford; 2004.
- Powderham AJ, Nicholson DP. The way forward. In: Nicholson DP, editor. *The observational method in geotechnical engineering*. London: Thomas Telford; 1996. p. 195–204.
- Powderham AJ. The observational method—learning from projects. *Proc Inst Civ Eng Geotech* 2002;155:59–69.
- Spross J, Johansson F, Stille H, Larsson S. Towards an improved observational method. In: Alejano LR, Peruchó A, Olalla C, Jiménez R, editors. *EUROCK 2014*. London: Taylor & Francis; 2014. p. 1435–40.
- Spross J, Johansson F, Uotinen LKT, Rafi JY. Using observational method to manage safety aspects of remedial grouting of concrete dam foundations. *Geotech Geol Eng* 2016;34:1613–30.
- Spross J, Johansson F. When is the observational method in geotechnical engineering favourable? *Struct Saf* 2017;66:17–26.
- Wallin S. Chasing a definition of “alarm”. *J Netw Syst Manage* 2009;17:457–81.
- Breznitz S. *Cry wolf: the psychology of false alarms*. Hillsdale: Lawrence Erlbaum Associates; 1984.
- Sättele M, Bründl M, Straub D. Reliability and effectiveness of early warning systems for natural hazards: concept and application to debris flow warning. *Reliab Eng Syst Saf* 2015;142:192–202.
- Sättele M, Krautblatter M, Bründl M, Straub D. Forecasting rock slope failure: how reliable and effective are warning systems? *Landslides* 2016;13:737–50.
- Paté-Cornell ME, Benito-Claudio CP. Warning systems: response models and optimization. In: Covelto VT, Lave LB, Moghissi A, Uppuluri VRR, editors. *Uncertainty in risk assessment, risk management, and decision-making*. New York: Plenum Press; 1987. p. 457–68.
- Olsson L, Stille H. Observation systems with alarm thresholds and their use in designing underground facilities. Stockholm: SKB; 2002.
- Der Kiureghian A, Ditlevsen O. Aleatory or epistemic? Does it matter? *Struct Saf* 2009;31:105–12.
- Spross J, Larsson S. On the observational method for groundwater control in the Northern Link tunnel project, Stockholm Sweden. *Bull Eng Geol Environ* 2014;73:401–8.
- Prästings A, Müller R, Larsson S. The observational method applied to a high embankment founded on sulphide clay. *Eng Geol* 2014;181:112–23.
- Wu TH. 2008 Peck lecture: the observational method: case history and models. *J Geotech Geoenviron Eng* 2011;137:862–73.
- Vrouwenvelder ACWM. Developments towards full probabilistic design codes. *Struct Saf* 2002;24:417–32.
- Straub D. Reliability updating with equality information. *Prob Eng Mech* 2011;26:254–8.
- Straub D. Value of information analysis with structural reliability methods. *Struct Saf* 2014;49:75–85.
- Ditlevsen O, Madsen HO. *Structural reliability methods*. In: Kgs. Lyngbyeditor. Coastal, maritime and structural engineering. Department of Mechanical Engineering, Technical University of Denmark; 2007.
- Hall WB. Reliability of service-proven structures. *J Struct Eng* 1988;114:608–24.
- Au S-K, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. *Prob Eng Mech* 2001;16:263–77.
- Koutsourelakis PS, Pradlwarter HJ, Schuëller GI. Reliability of structures in high dimensions, part I: algorithms and applications. *Prob Eng Mech* 2004;19:409–17.
- Bucher C. Asymptotic sampling for high-dimensional reliability analysis. *Prob Eng Mech* 2009;24:504–10.
- Zienkiewicz OC, Taylor RL, Zhu JZ. *The finite element method: its basis and fundamentals*. 7th ed. Oxford: Butterworth-Heinemann; 2013.
- Straub D, Papaioannou I, Betz W. Bayesian analysis of rare events. *J Comput Phys* 2016;314:538–56.
- Au S-K, Wang Y. *Engineering risk assessment with subset simulation*. Singapore: John Wiley & Sons; 2014.
- Papaioannou I, Betz W, Zwirgmaier K, Straub D. MCMC algorithms for subset simulation. *Prob Eng Mech* 2015;41:89–103.
- Schneider R, Thöns S, Straub D. Reliability analysis and updating of deteriorating systems with subset simulation. *Struct Saf* 2017;64:20–36.
- Zuev KM, Beck JL, Au S-K, Katafygiotis LS. Bayesian post-processor and other enhancements of subset simulation for estimating failure probabilities in high dimensions. *Comput Struct* 2012;92:283–96.
- Straub D, Papaioannou I. Bayesian updating with structural reliability methods. *J*

- Eng Mech 2014;141:04014134.
- [52] Au S-K. On MCMC algorithm for subset simulation. *Prob Eng Mech* 2016;43:117–20.
- [53] Comsol. COMSOL Multiphysics® version 5.3. Stockholm: Comsol AB; 2017.
- [54] Fenton GA, Naghibi F, Griffiths DV. On a unified theory for reliability-based geotechnical design. *Comput Geotech* 2016;78:110–22.
- [55] MathWorks. Matlab R2013b. Natick: MathWorks; 2013.
- [56] Gambino GF, Harrison JP. Multiple modes of rock slope instability – a limit state design approach. In: Schubert W, Kluckner A, editors. *Proceedings of the workshop design practices for the 21st century at EUROCK 2015 & 64th geomechanics colloquium*. Salzburg: Österreichische Gesellschaft für Geomechanik; 2015. p. 11–6.
- [57] Lü Q, Chan CL, Low BK. System reliability assessment for a rock tunnel with multiple failure modes. *Rock Mech Rock Eng* 2013;46:821–33.
- [58] Palmström A, Stille H. Ground behaviour and rock engineering tools for underground excavations. *Tunn Undergr Sp Tech* 2007;22:363–76.
- [59] Stanton N. Modelling human alarm initiated activities: implications for alarm system design. *IEE Colloquium on Man-Machine Interfaces for Instrumentation*. London: IEE; 1995. 8/1-8/4.