PRAKTISK ANVÄNDNING AV OBSERVATIONSMETODEN I ETT SANNOLIKHETSBASERAT RAMVERK:

Rekommendationer och beräkningsexempel

Johan Spross

Tobias Gasch

Fredrik Johansson

PRAKTISK ANVÄNDNING AV OBSERVATIONSMETODEN I ETT SANNOLIKHETSBASERAT RAMVERK:
Rekommendationer och beräkningsexempel

Practical use of the Observational Method in a reliability framework:
Recommendations and calculation examples

Johan Spross, KTH

Tobias Gasch, COMSOL

Fredrik Johansson, KTH

FÖRORD

Observationsmetoden framhålls ofta som ett viktigt verktyg för dimensionering av berg-tunnlar för att hantera de stora geologiska osäkerheter som man normalt ställs inför. Det saknas dock riktlinjer och rekommendationer för hur detta ska gå till i praktiken. Detta forskningsprojekt har därför syftat till att visa på hur observationsmetoden kan användas inom tunnelbyggnad i ett sannolikhetsbaserat ramverk. Bland de utvecklade verktygen finns en algoritm för fastställande av sannolikhetsbaserade larmgränser, som har en avgörande roll inom användningen av observationsmetoden.

Genom att använda det sannolikhetsbaserade ramverket kan man få värdefulla insikter om hur de rådande osäkerheterna kan påverka tunnelns säkerhet, på ett djupare sätt än vad som uppnås med deterministiska verktyg. Detta kan ge beslutsfattare en bättre för-ståelse av de risker man står inför, vilket bör ge ett mer kostnadseffektivt nyttjande av tillgängliga resurser i val av designlösning.

Stockholm 2022

Patrik Vidstrand

ii

PREFACE

The observational method is highlighted commonly as an important tool for the structural design of tunnels in hard rock to handle the large geological and geotechnical uncertainty that is present normally. There is however a lack of guidelines and recommendations for how to use this tool in present tunnel engineering practice. This research project has therefore aimed to address this issue by showing how the reliability-based framework for the observational method can be implemented into the design of rock tunnels. The developed tools include an algorithm for establishment of reliability-based alarm thresholds, which play a key role in the observational method.

Application of the reliability-based framework may provide valuable insights regarding the effect of geological and geotechnical uncertainties on the structural integrity of tunnels that are not captured by deterministic design methods. This can serve as additional input to decision makers understanding of the present risks in the project and thereby facilitate more well-informed decisions leading to more cost-effective tunnel engineering projects.

Stockholm 2022

Patrik Vidstrand

SAMMANFATTNING

Observationsmetoden framhålls ofta som ett viktigt verktyg för dimensionering av tunnlar i hårt berg. Genom att observera konstruktionens beteende under byggtiden kan man minska den så kallade epistemiska osäkerheten, som är kopplad till bristen på kunskap om de verkliga geotekniska förutsättningarna. Detta medför att man kan tillåta mindre konservativa designlösningar än vad som annars skulle krävas ifall man inte beaktade sådana observationer. Tack vare de senaste årens utveckling av nya sannolikhets-baserade verktyg och beräkningsmetoder, är det idag möjligt att använda observations-metoden i ett sannolikhetsbaserat ramverk. Därmed kan mätdata användas för att formellt uppdatera den beräknade brottsannolikheten för konstruktionen.

Det saknas dock idag både riktlinjer, rekommendationer och praktiska exempel på hur detta kan utföras i verkliga projekt. Detta forskningsprojekt har därför syftat till att visa hur det sannolikhetsbaserade ramverket för observationsmetoden kan användas vid dimensionering av bergtunnlar. En central del av observationsmetoden utgörs av hur man ska fastställa de larmgränser som anger oacceptabelt beteende hos konstruktionen och vars överskridande innebär att åtgärder behöver sättas in för att säkra konstruktionen.

Denna slutrapport ger en överblick över den konceptuella idén för det sannolikhets-baserade ramverket för observationsmetoden (då teorin presenteras i detalj i Spross doktorsavhandling som publicerats som BeFo-rapport 163). En utredning ges därefter av hur sannolikhetsbaserade larmgränser kan fastställas inom ramen för detta. Den teoretiska bakgrunden och praktiska tillämpningen utvecklas i detalj i två bifogade vetenskapliga artiklar. Artikel I är publicerad i Structural Safety och presenterar en beräkningsalgoritm för fastställandet av sannolikhetsbaserade larmgränser. Algoritmen använder sig av delmängdssimulering (eng. Subset Simulation) och fungerar med både analytiska och numeriska (finita element-)modeller i utvärderingen av den beaktade gränsfunktionen. Artikel II presenterar ett mer omfattande beräkningsexempel som inspirerats av de geologiska och geotekniska förhållandena vid Mälarpassagen i Förbifart Stockholms bergtunnlar. Exemplet visar hur larmgränser med hjälp av algoritmen kan fastställas för sprutbetongförstärkningens tillåtna deformation. Baserat på forskningsresultaten ger författarna rekommendationer avseende dels användningen av algoritmen, dels andra utmaningar som kan uppstå vid sannolikhetsbaserad dimensionering av bergtunnlar.

Sammanfattningsvis anser författarna att de utvecklade sannolikhetsbaserade verktygen kan vara särskilt användbara vid dimensionering av tunnlar i komplicerade geologiska förhållanden (t.ex. svagt berg och liten bergtäckning kombinerat med stora risker).

Användningen av verktygen kan då bidra med värdefulla insikter om hur de rådande osäkerheterna om de geotekniska förutsättningarna kan påverka konstruktionens säkerhet, som deterministiska verktyg inte förmår visa. Detta kan ge beslutsfattare en bättre förståelse av de risker man står inför, vilket bör ge ett mer kostnadseffektivt nyttjande av tillgängliga resurser i val av designlösning.

Nyckelord: Observationsmetoden; Larmgräns; Sannolikhetsbaserad dimensionering; Mätning; Sprutbetong; Säkerhet

SUMMARY

The observational method is highlighted commonly as an important tool for the structural design of tunnels in hard rock: by observing the structural behaviour during construction, a large portion of the epistemic uncertainty (lack of knowledge) that was present in the design phase can be eradicated. This allows for less conservative design solutions, which in many cases can lead to more economic designs. Following the development of new reliability-based tools and calculation procedures, it is today possible to apply the observational method within a reliability framework. This means that measurement data can be used to formally update the probability of unsatisfactory structural behaviour.

There is however a lack of guidelines and recommendations for how to use these new tools in practice. This research project has therefore aimed to address this issue by showing how the reliability-based framework for the observational method can be implemented into the design of rock tunnels. A key issue for the observational method is the establishment of thresholds toward unacceptable structural behaviour, as the violation of such a thresholds requires that prepared safety-enhancing design modifications are put into operation.

This final report outlines the conceptual idea of the reliability framework for the observational method (as it was previously published in Spross Ph.D. thesis from 2016) and discusses in detail how reliability-based alarm thresholds can be established. In two appended journal articles, the scientific basis for the procedure is elaborated in detail. Article I is published in Structural Safety and presents a computational algorithm for establishment of reliability-based alarm thresholds. The algorithm employs Subset simulation and can be used with both analytical and numerical (finite element) models in the evaluation of the considered limit state function. Article II presents a comprehensive calculation example inspired by the ground conditions present at the underwater passage of Lake Mälaren of the Stockholm Bypass rock tunnel project. The example shows how alarm thresholds can be established, using the algorithm, for the displacement of the shotcrete lining, as an indication of its structural integrity. Based on the findings of the research, recommendations are issued in this report regarding the practical implementation of the developed algorithm, as well as other challenging issues in reliability-based design of rock tunnels.

In conclusion, the developed tools are believed to be particularly valuable to design of tunnels in challenging geological conditions (e.g. weak rock with limited overburden and high risks). Their application may provide valuable insights regarding the effect of geological and geotechnical uncertainties on the structural integrity that are not captured

by deterministic design methods. This can serve as additional input to decision makers understanding of the present risks in the project and thereby facilitate more well-informed decisions leading to more cost-effective tunnel engineering projects.

LISTA ÖVER BIFOGADE ARTIKLAR

Inom ramen för detta forskningsprojekt har följande artiklar publicerats. Artiklarna återpubliceras i sin helhet i denna rapport med villkor enligt CC-BY 4.0 Creative Commons licens: https://creativecommons.org/licenses/by/4.0/

Artikel I

Spross, J. & Gasch, T. 2019. Reliability-based alarm thresholds for structures analysed with the finite element method. *Structural Safety, 76*, 174-183. https://doi.org/10.1016/j.strusafe.2018.09.004

Artikel II

Spross, J., Gasch, T. & Johansson, F. 2022. Implementation of reliability-based thresholds to excavation of shotcrete-supported rock tunnels, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards.* Accepterad för publicering 11 January 2022. https://doi.org/10.1080/17499518.2022.2046789

x

INNEHÅLL

xii

1. INLEDNING

Dimensionering av undermarksanläggningar såsom tunnlar och bergrum innebär att man säkerställer att konstruktionen byggs med tillräckligt stora säkerhetsmarginaler för de dimensioneringsfrågor som man har identifierat som viktiga. En stor utmaning ligger i att hantera den rådande osäkerheten kring bergets faktiska egenskaper och spännings-tillstånd. Även om man normalt utför undersökningar av borrkärnor och ibland även spänningsmätningar, så måste dimensioneringsarbetet ändå utföras under stor så kallad epistemisk osäkerhet, vilket innebär att det finns bristande kunskap om de rådande geologiska och geotekniska förhållandena. Detta gör att det är svårt att förutsäga vilket geotekniskt beteende som kan förväntas hos den tänkta konstruktionen.

Denna osäkerhet kan hanteras på olika sätt. Rent principiellt är det förstås möjligt att bygga undermarksanläggningar även med mycket liten kunskap om de rådande förhål-landena, men då krävs mycket stora säkerhetsmarginaler för att täcka in alla tänkbara förhållanden som skulle kunna uppträda. En sådan dimensioneringsstrategi anses oftast orimligt dyr. Inom undermarksbyggande har man därför kommit att förlita sig på den dimensioneringsmetod som går under namnet observationsmetoden.

Kortfattat, och något förenklat, innebär detta att man tar fram flera designlösningar, där en initial (även kallad preliminär) lösning passar för mer troliga förhållanden, medan övriga lösningar är anpassade efter andra, möjliga, men mer osannolika förhållanden. Samtidigt tar man fram dels planer för mätningar under byggtiden, dels åtgärder som gör det möjligt att byta från den initiala (preliminära) lösningen till någon av de andra. I vissa fall, exempelvis när tunnelförstärkning bestäms med hjälp av berg- och förstärk-ningsklasser, blir den "initiala designlösningen" i praktiken bara en av flera likställda typlösningar, där den mest lämpliga väljs baserat på mätresultatet.

Tanken med denna dimensioneringsstrategi är att man kan anpassa konstruktionen efter de faktiska förhållandena, istället för att använda mycket stora säkerhetsmarginaler. Därmed bör man kunna uppnå en ekonomiskt fördelaktig teknisk lösning. Att använda observationsmetodens principer för bergbyggande rekommenderas bland annat av Trafikverket i skriften "Projektering av bergkonstruktioner" (Trafikverket 2019). Obser-vationsmetoden har också accepterats i designkoder såsom Eurokod 7 (CEN 2004) som tillåten metod för att verifiera gränstillstånd i dimensioneringsarbetet. Eurokod 7 speci-ficerar därför ett antal krav på vad som behöver uppfyllas när metoden används.

Det finns dock ett antal tekniska svårigheter att hantera, när man ska uppfylla kraven som ställs på användandet av observationsmetoden. Dessa svårigheter undersöktes i detalj i Spross (2016) doktorsavhandling, som även finns utgiven som BeFo-rapport nr.

163. En sammanfattning av identifierade svårigheter och relevanta referenser ges nedan. Spross ger även förslag på en väg framåt för hur man kan hantera svårigheterna i praktiskt dimensioneringsarbete: genom att använda sig av sannolikhetsbaserade beräkningsmetoder kan man på ett metodiskt och förhållandevis enkelt sätt uppfylla de krav som ställs i Eurokod 7 observationsmetoden. Poängen med att använda sannolikhetsbaserade beräkningsmetoder är att man kan hantera osäkerheten i dimensioneringsarbetet på ett stringent sätt genom att ansätta sannolikhetsfördelningar för de osäkerheter som finns avseende de geotekniska förhållandena. Sannolikheten för kombinationer av flera ogynnsamma förhållanden samtidigt kan därmed analyseras mer på djupet, än vad en klassisk sensitivitetsanalys ger.

I en vetenskaplig artikel (Spross & Johansson 2017), som ingick i avhandlingen, visas hur observationsmetoden kan tolkas i ett sannolikhetsbaserat ramverk. Bland annat visas hur man rent principiellt kan ta fram larmgränser för acceptabelt beteende baserat på en i förväg fastställd acceptabel brottsannolikhet för det gränstillstånd som man dimensionerar mot. Spross doktorsarbete fokuserade dock i huvudsak på de underliggande principerna för det sannolikhetsbaserade ramverket och presenterade endast förenklade och begränsade beräkningsexempel för att illustrera dess tillämpning.

1.1   Syfte och omfattning

Detta forskningsprojekt har syftat till att göra det sannolikhetsbaserade ramverket för observationsmetoden mer tillgängligt för den praktiserande ingenjören. I denna slutrapport ges därför rekommendationer och fördjupade diskussioner kring hur ramverket kan användas i mer avancerade dimensioneringsfall inom undermarksbyggande. Stor vikt har i detta forskningsprojekt lagts vid att göra ramverket användbart tillsammans med beräkningstunga numeriska analysverktyg såsom finita elementmetoden (FEM), eftersom sådana verktyg är vanligt förekommande i praktiskt dimensioneringsarbete i dagens undermarksbyggande. Hur detta låter sig göras i praktiken visas i ett mer omfattande tillämpningsexempel.

Läsaren av den svenska delen av rapporten förutsätts ha grundläggande kunskaper i statistik och principer för dimensionering i nivå med vad som undervisas på våra tekniska högskolors civilingenjörsutbildningar inom anläggningsbyggande. I de bifogade artiklarna används i vissa fall en mer avancerad terminologi och matematisk formulering, även om koncepten är desamma som de som beskrivs i den svenska huvudtexten.

1.2   Översikt över innehållet

Denna slutrapport består av en huvudtext på svenska och två bifogade vetenskapliga artiklar på engelska. Huvudtexten redogör för de övergripande principerna för det

sannolikhetsbaserade ramverket och ger generella rekommendationer för dess tillämpning samt en diskussion kring möjliga utmaningar som kan uppstå.

Artikel I presenterar en beräkningsalgoritm för framtagande av sannolikhetsbaserade larmgränser för oacceptabelt beteende. Algoritmen är utvecklad för att vara effektiv nog att kunna hantera relativt tunga numeriska beräkningsmodeller.

Artikel II presenterar ett mer omfattande tillämpningsexempel som är inspirerat av förhållandena vid Mälarpassagen i byggandet av Förbifart i Stockholm. I exemplet visar vi hur man kan ta fram sannolikhetsbaserade larmgränser för sprutbetongförstärkningens deformation. Sådana larmgränser är en central del i de mät- och kontrollprogram som behövs vid användande av observationsmetoden. Erfarenheten från det utförda forskningsprojektet visar att framtagandet av sannolikhetsbaserade larmgränser är den svåraste delen rent beräkningstekniskt.

2.   DIMENSIONERING MED OBSERVATIONSMETODEN INOM BERGBYGGANDE

2.1   Observationsmetoden – grundläggande principer

Observationsmetoden sägs ofta ha utvecklats av Terzaghi och Peck, med referens till Pecks (1969) artikel i Geotechnique som för första gången formulerade observationsmetodens principer i vetenskapliga sammanhang. Dessa principer var dock Pecks försök att generalisera den, enligt Peck, mycket framgångsrika ingenjörsmetodik som Terzaghi använde sig av under tiden som praktiserande ingenjör.

Även om Peck marknadsförde observationsmetoden som ett nytt alternativ för geoteknisk dimensionering, så är principerna i stora drag i linje med det som i århundraden varit klassiskt, gott ingenjörsarbete. Det finns till och med vissa historiska belägg för att liknande principer användes av fenicier vid ett kanalbygge redan före Kristi födelse (Herodotus ca 430 f.Kr.). Även i Sverige har observationsmetodens övergripande idé varit känd länge. Statens Järnvägars geotekniska kommission (1922) föreslog exempelvis övervakning av jordrörelser tillsammans med varningssystem som ett kostnadseffektivt sätt att undvika olyckor, eftersom man insåg att garanterat säkra järnvägsbankar helt enkelt inte skulle vara ekonomiskt försvarbart. Samtidigt: Pecks och Terzaghis arbete med att marknadsföra observationsmetoden var välbehövligt med tanke på det stora risktagande som de ofta noterade i geotekniska projekt på den tiden.

Huvuddragen i observationsmetodens idé presenterades i rapportens inledning. För att fungera som en alternativ metod för att verifiera gränstillstånd i en designkod fanns dock ett behov att skriva om Pecks beskrivna arbetsgång till en uppsättning skall-krav. Detta gjordes i samband med skapandet av Eurokod 7, vars krav avseende observationsmetodens användning citeras i det följande. Den som är intresserad av den historiska utvecklingen kan gärna jämföra med Pecks (1969) ursprungliga formulering, som även diskuteras i Spross (2016). I Eurokod 7 anger "P" en så kallad princip som det inte får göras avsteg ifrån.

> *(1) När förutsägelsen av det geotekniska beteendet är svår kan det vara lämpligt att tillämpa den metod som benämns 'observationsmetoden', där dimensioneringen följs upp under byggnadsskedet.*

> *(2)P Följande krav skall uppfyllas innan utförandet påbörjas:*
> - *Acceptabla gränser för beteendet skall bestämmas;*

- *Gränserna för möjligt beteende skall beräknas och det skall visas att sannolikheten för att det verkliga beteendet ligger inom de acceptabla gränserna är godtagbar;*
- *En plan för uppföljning skall tas fram som skall visa om det verkliga beteendet ligger inom acceptabla gränser. Uppföljningen skall på ett tillräckligt tidigt stadium klargöra detta och med tillräckligt korta tidsintervall för att framgångsrikt kunna vidta korrigerande åtgärder* [sic]*;*
- *Responstiden hos mätinstrumenten och i sättet att analysera resultaten skall vara tillräckligt snabbt för att möjliggöra förändringar i systemet;*
- *En plan för korrigerande åtgärder skall upprättas, vilken kan följas om uppföljningen visar ett beteende som ligger utanför acceptabla gränser.*

(3)P Under byggnadsskedet skall uppföljningen utföras som planerat.

(4)P Resultaten av uppföljningen skall utvärderas vid lämpliga steg och de planerade korrigerande åtgärderna skall vidtas om gränserna för beteendet överskrids.

(5)P Uppföljande instrumentering skall antingen bytas ut eller utökas om den inte ger tillförlitliga data av avsett slag eller i tillräcklig mängd.

I sin analys av Eurokodens definition av observationsmetoden noterar Spross (2016) att definitionen är mycket generell, men att detta kan ha sin förklaring i att den ska vara tillämplig i väldigt många olika typer av geotekniska byggprojekt. De riktlinjer för tillämpning som finns tillgängliga (Frank et al. 2004) ger däremot förvånansvärt få detaljer och råd kring hur Eurokodens krav på observationsmetoden ska tolkas i praktiken. Här i ligger själva grunden till behovet av den forskning om observationsmetoden som föreliggande rapport är en del av: Att ge tydligare och mer detaljerade råd kring hur observationsmetoden ska användas för att säkerställa att konstruktionen uppfyller ställda säkerhetskrav på ett ekonomiskt sätt. Detta gäller särskilt:

1) hur man ska kunna *visa* att det är en godtagbar sannolikhet att det verkliga beteendet ligger inom de acceptabla gränserna (d.v.s. att den initiala designen med godtagbar sannolikhet kommer att lyckas utan att dyra åtgärder behöver sättas in),

2) hur ordet *godtagbar* ska definieras i detta sammanhang,

3) hur larmgränser ska sättas så att konstruktionen är tillräckligt säker utan att bli överdimensionerad.

Svaren på dessa tre frågor kommer vi att återkomma till i efterföljande kapitel.

2.2 Erfarenheter av observationsmetoden inom bergbyggande

Sverige har en relativt lång tradition av att använda observationsmetoden och liknande principer inom bergbyggande. I Sverige gick dock detta länge under namnet "Aktiv design", vilket först beskrevs vetenskapligt av (Stille 1986). Flera forskningsprojekt har även utförts i Sverige (med finansiering av BeFo) i syfte att förbättra och sprida användandet av observationsmetoden inom svenskt bergbyggande. Holmberg och Stille (2007, 2009) diskuterade exempelvis hur observationsmetoden kan användas tillsammans med deformationsmätningar vid tunnelbyggande, vars koncept senare har utvecklats i Bjureland et al. (2017) och i Artikel II som bifogats denna rapport. Zetterlund (2014) diskuterade observationsmetodens roll vid dimensionering av injektering. Olika möjliga tillämpningsområden för observationsmetoden har även studerats inom ramen för Spross doktorsprojekt (t.ex. Spross & Larsson 2014, Spross et al. 2016).

Stille & Holmberg (2010) menar att svårigheterna med att förutsäga det geotekniska beteendet vid byggande av undermarksanläggningar i hårt, sprickigt berg i stort kan hänföras till tre kategorier:

1. Osäkerheter i lägesbestämning av förekommande bergkvaliteter samt deras faktiska variation,

2. Osäkerhet kring samverkan mellan bergmassa och installerad förstärkning,

3. Osäkerhet kring kvaliteten på den utförda förstärkningen.

Dessa tre kategorier kan i sig exemplifieras med följande tre situationer där observationsmetoden kan vara tillämpbar, enligt Stille & Holmberg (2010):

1. Kartering av tunnelväggen, d.v.s. observation av faktisk bergkvalitet, kan användas för att anpassa förstärkningen till den faktiska bergkvaliteten (istället för att endast förlita sig på prognosen från förundersökningen).

2. Mätning av tunneldeformation för att bedöma lokala stabilitetsproblem i enskilda tunnelsektioner.

3. Mätning (kontroll) av sprutbetongens tjocklek för att verifiera att utfört arbetet uppfylla ställda (säkerhets)krav.

Holmberg & Stille (2007) identifierar några centrala aspekter avseende byggprojektets organisation, för att observationsmetoden ska kunna användas framgångsrikt:

- Observationsmetoden ska vara integrerad i dimensioneringsprocessen redan från början.

- Val av mätmetod ska göras baserat på en noggrann analys av vilka dimensioneringsfrågeställningar som är kritiska i det aktuella fallet.

- Observationsmetoden ska ha en tydlig roll under byggtiden, vara integrerad i produktionen, samt ha stöd i kontrakts- och ersättningsformen.

Den sistnämnda punkten, avseende avsaknad av lämpliga kontrakts- och ersättningsformer, har länge ansetts vara en bromskloss i tillämpningen av observationsmetoden i Sverige (se t.ex. Kadefors & Bröchner 2008). Kontraktuella aspekter på observationsmetodens tillämpning diskuteras dock inte vidare i denna rapport.

Internationellt sett är litteraturen om användning av observationsmetoden inom bergbyggande mycket begränsad, men Maidl et al. (2011) och Miranda et al. (2015) utgör två undantag. Observationsmetodens principiellt viktiga roll för bergbyggande har dock särskilt lyfts fram internationellt i artiklarna av Spross et al. (2018, 2020), vilka diskuterar hur en designkod bör vara uppbyggd för att fungera inom bergbyggande. En övergripande diskussion av dessa frågor ges på svenska i tillhörande BeFo-rapport 190 (Spross et al. 2019).

För en genomgång av tillgänglig litteratur om tillämpning av observationsmetoden inom andra delar av geotekniken än bergbyggande hänvisar vi till Spross (2016) doktorsavhandling.

3. ÖVERSIKT ÖVER DET SANNOLIKHETSBASERADE RAMVERKET FÖR OBSERVATIONSMETODEN

Det sannolikhetsbaserade ramverket för observationsmetoden har tidigare översiktligt beskrivits på svenska i en artikel till Bergmekanikdagen av Spross (2017). Nedan återfinns en reviderad och delvis utökad version av den beskrivningen.

3.1 Några centrala begrepp

3.1.1 Gränstillstånd och tillåten brottsannolikhet

Innan vi presenterar själva ramverket behöver ett antal centrala begrepp definieras. Sannolikhetsbaserad dimensionering förutsätter att den beaktade dimensionerings-frågeställningen kan beskrivas matematiskt med en gränsfunktion som definierar gräns-tillståndet mellan acceptabelt och oacceptabelt beteende hos konstruktionen. Detta skrivs ofta som

$$G(\boldsymbol{X}) = 0, \tag{1}$$

där G betecknar en funktion och $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]$ är en vektor som samlar de påverkande stokastiska variablerna och som alltså matematiskt beskriver de geotekniska osäkerheterna. Sannolikhetsbaserad dimensionering går ut på att säkerställa att sannolikheten att gränstillståndet uppnås, d.v.s. $G(\boldsymbol{X}) \leq 0$, är mindre än en viss tillåten brott-sannolikhet hos konstruktionen, p_{FT}. Notera att både brottgränstillstånd och bruksgräns-tillstånd hanteras på samma sätt, men med olika tillåtna p_{FT}. Vilka p_{FT} som kan tillåtas brukar normalt framgå av tillämplig dimensioneringsstandard. Vi har alltså att uppfylla att den beräknade brottsannolikheten p_F är

$$p_F = P[G(\boldsymbol{X}) \leq 0] \leq p_{FT}. \tag{2}$$

Notera även att en del dimensioneringsfrågor inom bergbyggande kan vara svåra eller olämpliga att beskriva matematiskt som gränstillstånd. Detta inkluderar exempelvis nedbrytning av injekteringsbruk med tiden eller problem med bildning av istappar i tunneltaket. Sådana frågeställningar kan inte hanteras inom detta ramverk.

3.2 Poängen med ramverket

I det sannolikhetsbaserade ramverket kombineras sannolikhetsbaserad dimensionering med bayesiansk beslutsteori (mer om detta nedan). Detta upplägg gör det möjligt att jämföra den förväntade kostnaden för olika möjliga designalternativ för en konstruktion, samtidigt som ramverket säkerställer att samhällets krav på acceptabel säkerhet hos

konstruktionen uppnås, oavsett vilket designalternativ som i slutändan väljs. Den observanta läsaren har kanske noterat att Eurokod 7 faktiskt inte tydliggör vilken säkerhetsnivå ska uppnås hos den färdiga konstruktionen när observationsmetoden används, vilket kan tyckas märkligt. (Den välvilliga läsaren av Eurokoden kan förstås välja att tolka in att det är upp till ingenjören att välja en metod att verifiera konstruktionens säkerhet, utöver att tillämpa observationsmetodens principer, men detta framgår alltså inte explicit i dagens version av Eurokod 7).

Ramverket kommer dessutom åt stötestenarna i tillämpningen av Eurokod 7 som nämndes ovan: hur man ska sätta larmgränser och hur man därefter visar att sannolikheten att dessa överskrids är godtagbar. I det följande ges en konceptuell beskrivning av hur detta går till. En komplett beskrivning av teorin finns i Spross & Johansson (2017). Artikeln är fritt tillgänglig för nedladdning.

3.3 Så fungerar ramverket

3.3.1 Beslutsträdet

Med utgångspunkt i klassisk bayesiansk beslutsteori kan man dela upp tillämpningen av observationsmetoden i fyra delar (Figur 1): Dimensioneringsfas E, Mätning Z, Åtgärd A och Slutgiltigt beteende Θ. Från noden till vänster i figuren utgår alla tänkbara designalternativ. Det övre huvudalternativet är en möjlig tillämpning av observationsmetoden och det undre är en möjlig konventionell designlösning utan möjlighet till förändrad design. (I figuren jämförs två alternativ, men rent principiellt finns förstås ett stort antal möjliga utformningar och konstruktionsmetoder att ta ställning till. Dessa potentiella alternativ representeras av de korta streck som utgår från vissa noder).

Om man använder observationsmetoden ser vi att beroende på vilket mätresultat, Z, som vi får, så får vi olika slutlig design. Om vi får ett mätresultat som ligger under den fastställda larmgränsen (resultatet z_1 i figuren), så är den initiala (preliminära) designen acceptabel och ingen åtgärd behöver sättas in. För mätresultat som ligger över larmgränsen krävs däremot en åtgärd. I figuren visas det generella fallet, där man principiellt kan tänka sig att man har olika åtgärder beroende på hur "illa" mätresultatet är, d.v.s. man har flera larmgränser.

Den sista förgreningen, Θ, till höger i figuren, avser om gränstillståndet överskridits eller ej, alltså ifall gränsfunktionen $G(X) > 0$ (ej brott) eller ifall $G(X) < 0$ (brott). Benämningen "brott" används här för enkelhets skull; man kan lika gärna definiera dessa två situationer
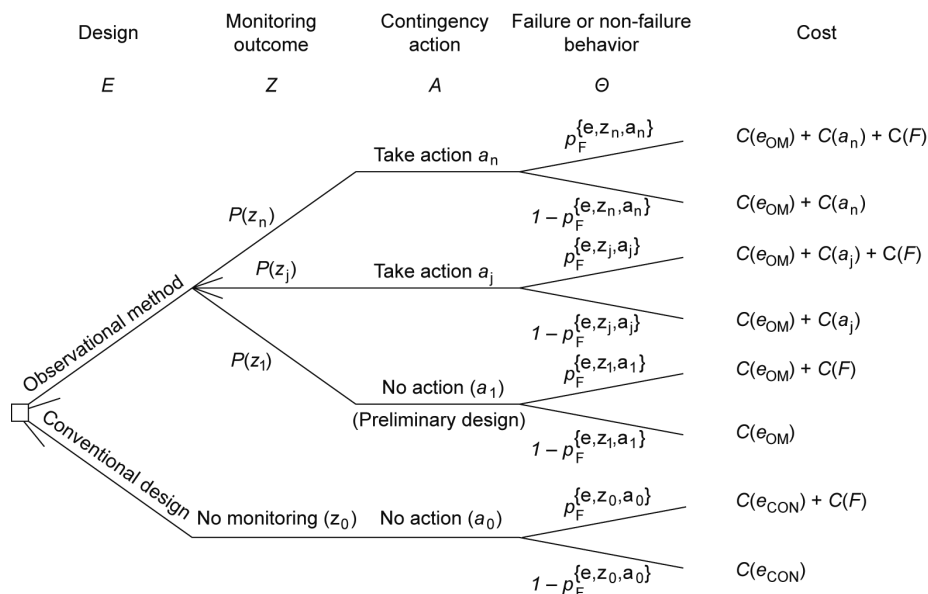
Figur 1. *Ramverket delar upp observationsmetoden i fyra delar: val av design, mätresultat, åtgärd och slutligt beteende. Varje möjlig design och utfall förknippas med en bedömd kostnad. Den förväntade kostnaden för att använda observationsmetoden ges av en sammanvägning av sannolikheterna för de olika utfallen. (Figur: Spross och Johansson 2017, CC-BY licens, http://creativecommons.org/licenses/by/4.0)*

som med hjälp av en bruksgräns eller någon annan gräns som man inte vill ska överskridas. Längst till höger redovisas vilka kostnader som kan hänföras till respektive utfall: kostnaden för den ursprungliga designlösningen, $C(e_{OM})$; kostnaden för eventuell åtgärd, $C(a)$; och kostnaden för eventuellt överskridet gränstillstånd, $C(F)$.

Med hjälp av sannolikhetsbaserade metoder kan man sedan dels beräkna sannolikheten för överskridet gränstillstånd, p_F, dels beräkna sannolikheten att larmgränsen överskrids eller underskrids, $P(z)$. Så som diskuteras i Spross & Johansson (2017) kommer den optimala designlösningen med observationsmetoden att utgöra en kompromiss mellan hur konservativ den initiala designen är och kostnaderna för eventuella åtgärder. Alltså: ju dyrare åtgärderna är desto mindre sannolikt får det vara att de behövs. Om åtgärderna är mycket dyra behöver därmed den initiala designen vara mer konservativ. Den optimala designlösningen, e_{opt}, (med detta avses utformningen av initial design inklusive förberedda mätplaner / kontrollprogram och planer för åtgärder) ges då av att tillämpa följande minimering på beslutsträdet i Figur 1:

$$e_{\text{opt}} = \arg \min_E \left\{ \sum_{j=1}^n P(z_j|e_i) C(e_i, z_j, a_j) \mid p_F^{\{e_i, z_j, a_j\}} \le p_{FT} \right\}, \tag{3}$$

där $P(z_j|e_i)$ är sannolikheten för att uppmäta mätresultatet z_j ifall designvariant e_i används och $C(e_i, z_j, a_j)$ är den förväntade kostnaden för respektive möjligt utfall (högra kolumnen i Figur 1). Denna kostnad kan beräknas genom en sammanvägning av kostnaderna $C(e_i, z_j, a_j, \theta_k)$ för de två möjliga utfallen av det slutliga beteendet hos konstruktionen (brott eller inte brott), med avseende på den beräknade sannolikheten för detta, d.v.s. $P(\theta_k|e_i, z_j, a_j)$, för den aktuella grenen i beslutsträdet:

$$C(e_i, z_j, a_j) = \sum_{k=1}^m C(e_i, z_j, a_j, \theta_k) P(\theta_k|e_i, z_j, a_j). \tag{4}$$

I det följande redovisas först de viktigaste underliggande antagandena som görs när man använder bayesiansk beslutsteori. Därefter redovisas hur de respektive komponenterna i ekvation (3) kan beräknas.

3.3.2 Underliggande antaganden vid användning av bayesiansk beslutsteori

Statistik delas ofta upp i två olika "skolor", beroende på vilket sätt som man tolkar beräkningsresultaten. De flesta lär sig det klassiska, frekventistiska tolkningssättet i kurser på gymnasiet och i högskolans grundkurser i statistik. Sannolikheten för ett visst utfall på ett framtida tärningskast kan beräknas genom att först observera utfallen av ett stort antal tidigare kast med samma tärning. Denna tolkning av statistik ger dock problem inom geotekniken. En beräkning av en brottsannolikhet skulle då kräva att ett stort antal exakt likadana konstruktioner byggs under exakt samma förhållanden, vilket naturligtvis blir orimligt.

Den bayesianska skolan, namngiven efter Bayes (1763), utgår istället från att sannolikheten är subjektiv och representerar användarens *grad av tilltro* (eng. *degree of belief*) avseende händelsen. Därmed representerar alltså en beräknad brottsannolikhet vad användarens uppfattning är om sannolikheten för brott i konstruktionen. Detta sker genom att man kombinerar kända data med andra informationskällor, såsom tillgänglig litteratur eller den egna expertbedömningen. Stokastiska variabler i den bayesianska skolan är alltså användarens bästa uppfattning om den statistiska fördelningsfunktionen i respektive variabel. Det bayesianska synsättet är därför mycket användbart när man som bergmekaniker statistiskt ska beskriva relevanta geotekniska parametrar, eftersom dessa ofta behöver beskrivas baserat på relativt lite objektivt inhämtad data från mätningar och andra undersökningar. Sådan så kallad sannolikhetsbaserad karaktärisering

av geotekniska parametrar är en mycket viktigt del av dimensioneringsarbetet, men då detta finns väl beskrivet i litteraturen (t.ex. Phoon & Kulhawy 1999a, Phoon & Kulhawy 1999b, Müller et al. 2014, van der Krogt et al. 2019) går vi inte djupare in på detta här.

Skillnaden mellan frekventistisk och bayesiansk statistik kan också uttryckas som att frekventisten anser att världen är full av okända konstanter som kan bestämmas exakt efter ett stort antal upprepade försök, medan bayesianen anser att världen är full av slumpvariabler som kan beskrivas med statistiska fördelningar. För en djupdykning i detta ämne rekommenderas artikeln av Vrouwenvelder (2002) och textboken av Baecher & Christian (2003). De statistiska beräkningar som utförs i denna rapport ska alltså tolkas enligt detta bayesianska synsätt.

En särskild gren av bayesiansk statistik utgörs av beslutsteori. Denna strävar efter att beräkna det ekonomiskt rationella beslutet givet problemets fastställda ramar. Utmärkta textböcker i ämnet är klassikerna av Raiffa & Schlaifer (1961) och Benjamin & Cornell (1970). (Vetenskapliga funderingar kring rationella beslut kan dock härledas tillbaka ända till Pascals trossats från 1670, enligt vilken man bör anta att Gud och himmelriket finns och därför leva gudfruktigt, eftersom vinsten med att göra detta är oändlig, medan de förlustelser som man avstår som gudfruktig har ett begränsat värde.)

I denna rapport utgår vi från att det mest rationella är att sträva efter maximerande av förväntad nytta, enligt von Neumann & Morgenstern (1944) hypotes. Detta har dock förenklats genom att anta att nytta kan beskrivas i monetära termer och att nytta är omvänt proportionell mot uppkomna kostnader, C, så att det optimala beslutet är det som minimerar den förväntade kostnaden, $E(C)$, så som beskrivs i ekvation (3). Med dessa antaganden sägs det rekommenderade beslutet vara riskneutralt (att jämföra med så kallade *risk-aversiva* –riskundvikande – beslut som favoriserar lösningar med små men säkra vinster eller förluster, framför lösningar med stora men osäkra vinster eller förluster).

3.3.3 Beräkning av brottsannolikhet givet uppmätt mätresultat

Brottsannolikheten för respektive gren räknas alltså givet uppmätt mätresultat Z och därefter införda åtgärder A. Operatorn "arg" betyder argument och innebär att ekvationen tar fram den designlösning som (i detta fall) minimerar kostnaden för konstruktionen bland samtliga tänkbara lösningar i vektorn E. Den införda begränsningen,

$p_{\mathrm{F}}^{\{e_i, z_j, a_j\}} \leq p_{\mathrm{FT}}$, i ekvation (3) säkerställer att brottsannolikheten är mindre än eller lika med den fastställda acceptabla brottsannolikheten för det aktuella gränstillståndet. Det gör att endast designlösningar som är acceptabla med avseende på tillåten brottsannolikhet kan föreslås av ekvationen.

Brottsannolikheten $p_{\mathrm{F}}^{\{e_i, z_j, a_j\}}$ kan beräknas med hjälp av vanlig Monte Carlo simulering:

$$p_{\mathrm{F}} = \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} I_{\mathrm{F}}(\mathbf{x}), \tag{5}$$

där $I_{\mathrm{F}}(\mathbf{x})$ är en så kallad indikatorfunktion för brott som ger $I_{\mathrm{F}}(\mathbf{x}) = 1$ ifall $G(\mathbf{x}) \leq 0$ och $I_{\mathrm{F}}(\mathbf{x}) = 0$ i övriga fall, och N_{MC} är antalet simuleringar av värden för slump-variablerna i $\mathbf{x}$. I Monte Carlo-simuleringen måste man dock beakta det man känner till om mätresultatet Z. Låt oss säga att vi observerar parameter X_1. Om vi beräknar brott-sannolikheten för fallet att initial designlösning är acceptabel (d.v.s. mätningen under-skrider larmgränsen), så måste fördelningen för X_1 trunkeras vid larmgränsen, så att de simulerade värdena i Monte Carlo simuleringen bara kan anta värden lägre än larm-gränsen. (Detta kan i praktiken enkelt lösas genom att sortera ut otillåtna simulerade värden med en *if*-sats). Det omvända gäller förstås för fallet att larmgränsen överskridits och åtgärd satts in: då används endast värden som överskrider larmgränsen. Simulerade värden efter trunkering för en normalfördelning visas som exempel i Figur 2. (Hur själva larmgränsen x_{alarm} fastställs redogörs för i nästkommande avsnitt.)

Vanlig Monte Carlo simulering kan dock kräva många utvärderingar av den aktuella gränsfunktionen G. Ifall en analytisk gränsfunktion används, så är antalet utvärderingar N sällan något praktiskt problem, även om antalet utvärderingar kan bli mycket stort ifall mycket låg brottsannolikhet eftersträvas hos konstruktionen. Om man däremot måste utvärdera en numerisk modell varje gång, exempelvis genom att köra en beräk-ning med finita element, så kan en vanlig Monte Carlo simulering bli mycket tids-krävande. Detta problem kan dock minskas genom att man använder effektivare simu-leringsmetoder, som minskar antalet
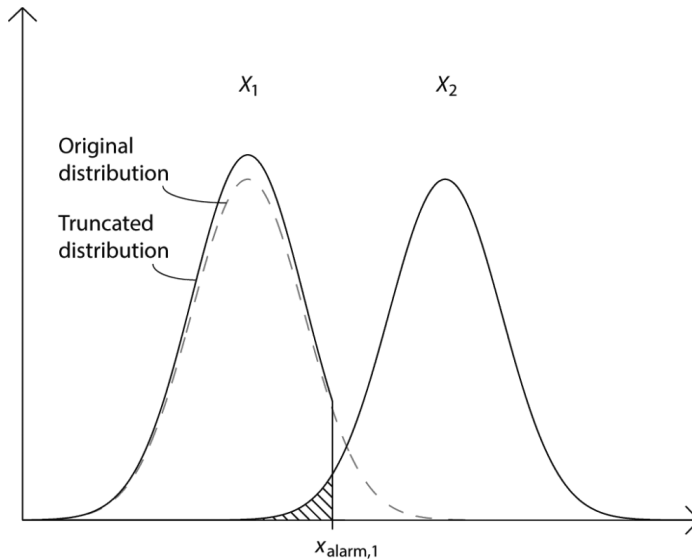
Figur 2. *Genom att trunkera variabel X_1 vid larmgränsen $x_{alarm,1}$ kan man säkerställa att brottsannolikheten hålls tillräckligt liten, för gränsfunktionen $G = X_2 - X_1 = 0$. Sannolikheten för kombinationen av farliga utfall där $X_2 < X_1$ blir då acceptabelt liten. (Figur: Spross och Johansson 2017)*

utvärderingar som krävs. I den bifogade Artikel I redogör vi för hur man kan använda så kallad Subset-simulering för detta ändamål. Grundprinciperna för Subset-simulering ges i Avsnitt 3.4.

3.3.4 Fastställande av larmgränser givet tillåten brottsannolikhet

För att kunna utföra trunkeringen som diskuteras i föregående avsnitt, behöver först larmgränsen fastställas. Konceptuellt sett görs detta såsom visas i Figur 2. Genom en iterativ beräkning prövar man för den observerade parametern x_1 helt enkelt olika tänkbara larmgränser $x_{alarm,1}$ tills man hittat den larmgräns som ger en trunkering, som medför att den beräknade brottsannolikheten för den analyserade gränsfunktionen G blir just den tillåtna brottsannolikheten p_{FT}. Detta beskrivs med följande ekvation:

$$P\big(G(\boldsymbol{X}) < 0 | x_1 \leq x_{alarm,1}\big) = p_{FT}, \tag{6}$$

som ska tolkas som att brottsannolikheten ska vara lika med p_{FT} så länge som respektive mätvärde x_1 för de observerade parametrarna underskrider sin larmgräns $x_{alarm,1}$.

I Artikel I beskrivs en algoritm för hur denna iterativa beräkning kan automatiseras och effektiviseras med hjälp av Subset-simulering. Denna algoritm kan i princip implementeras direkt som en modul eller ett tillägg i koden till finita-element-beräkningsprogram, vilket i så fall i framtiden skulle minska behovet av egen programmering för bergmekanikern.

Mer generellt kan man dock tänka sig att man mäter flera olika parametrar i vektorn X, vilket i så fall gör att en hel uppsättning tillhörande larmgränser x_{alarm} ska tas fram ur ekvation (6), så att vi får $P(G < 0 | x \leq x_{alarm}) = p_{FT}$. Eftersom vi då får flera okända larmgränser att bestämma ur ekvationen blir den underbestämd. Om det är värt besväret (kostnaden) att använda flera larmgränser, och vilka kombinationer av gränsvärden som då blir optimala, blir ett beslutsteoretiskt problem. Lösningen på detta beror bland annat av om den extra informationen är värd ökade kostnader för att ha flera olika mätsystem. Detta kan ses som olika varianter på designlösningar (d.v.s. olika mätplaner / kontrollprogram) och kan i princip analyseras genom att införa fler designalternativ i beslutsträdet, som sedan analyseras med ekvation (3).

I framtagandet av larmgränsen behöver man beakta eventuella mät- och transformationsfel i den tillämpade mätmetoden. Exempelvis kan man vilja göra "observationer" av överskriden hållfasthet baserat på mätningar av deformationer, vilket vi gör i Artikel II. Tillvägagångssättet för att hantera detta beskrivs i detalj i Rackwitz & Schrupp (1985). Det förväntade *mätresultatet Y* ses då som en egen stokastisk variabel:

$$Y = X\varepsilon, \tag{7}$$

där ε är det förväntade mätfelet, som här exempelvis kan vara normalfördelat med medelvärde lika med 1 och ha någon variationskoefficient som motsvarar mätfelets spridning. *Y* ersätter då X_i i begränsningen i ekvation (6). Mätfelets storlek är dock i dagsläget okänt för många mätmetoder inom bergbyggande, men bedöms kunna vara betydande för en del metoder. Avseende konvergensmätningar i bergtunnlar har Gothäll (2011) föreslagit en analysmetod som förbättrar databehandlingen och feluppskattningen, eftersom mätfelet under vissa förhållanden kan bli större än rörelserna man försöker mäta upp.

3.3.5 Sannolikheten för att det verkliga beteendet ligger inom larmgränserna

Såsom framgår av Eurokodens krav på observationsmetoden behöver man beräkna sannolikheten för att det verkliga beteendet ligger inom de acceptabla gränserna (d.v.s.

inom de fastställda larmgränserna). När man har fastställt larmgränsen enligt föregående avsnitt kan denna sannolikhet enkelt beräknas. Det är helt enkelt sannolikheten att de observerade mätresultaten överskrider larmgränsen:

$$P(X_1 > x_{\text{alarm},1}), \tag{8}$$

vilket enkelt beräknas med hjälp av vanlig Monte Carlo-simulering.

3.3.6 Är sannolikheten "godtagbar" för att verkligt beteende ligger inom gränserna?

Begreppet *godtagbar* är inte definierat i Eurokod 7, trots att detta är ett centralt begrepp i kraven avseende observationsmetoden:

> • *Gränserna för möjligt beteende skall beräknas och det skall visas att sannolikheten för att det verkliga beteendet ligger inom de acceptabla gränserna är* godtagbar

Publicerade riktlinjer för Eurokodens tillämpning (Frank et al. 2004) ger heller ingen ytterligare information. I Spross (2016) görs tolkningen att kravet på "godtagbar sannolikhet" avser det ingenjörsmässiga övervägandet om sannolikheten för att åtgärder behöver sättas in är värd de kostnader och eventuella förseningar som detta kan ge projektet. Alltså: om denna sannolikhet inte är godtagbar måste en annan designlösning än den föreslagna väljas.

En möjlig matematisk formulering av detta krav kan göras med hjälp av ekvation (3) och beslutsträdet i Figur 1. Om det efter utvärdering av förväntade kostnader för alla rimliga designlösningar med ekvation (3) visar sig att en (troligen mycket) konservativ designlösning endast baserad på förundersökningen, utan några som helst mätningar under byggtiden eller förberedda åtgärdsplaner, är den optimala lösningen e_{opt}, så är kravet på "godtagbar sannolikhet" inte uppfyllt för någon variant på designlösning med observationsmetoden. Den beslutsteoretiska beräkningen har då visat att eventuella möjliga åtgärdsplaner för de beaktade alternativen med observationsmetoden är för dyra för att tillåtas med någon sannolikhet. Alternativt är mätsystemet i sig alltför dyrt för att vara värt att installera. Om däremot e_{opt} beskriver en lösning med observationsmetoden, så är kravet på godtagbar sannolikhet implicit uppfyllt.

Det kan dock diskuteras i vilken mån denna princip i Eurokoden behöver uppfyllas. Även om den ovan beskrivna principen för att bedöma godtagbar sannolikhet, är möjlig att analysera i teorin, så kan det förstås i praktiken blir mycket tidskrävande att beskriva

samtliga möjliga designlösningar i sådan detalj att de kan analyseras med sannolikhets-baserade metoder. Det kan därför vara nödvändigt med principiella förenklingar, där ett fåtal huvudalternativ översiktligt analyseras i ett första skede.

Värd att notera är också diskussionen i Spross et al. (2016), i vilken det argumenteras för att man i vissa fall bör tillåtas att bortse från kravet på att bedöma om sannolikheten för att man kommer att ligga inom larmgränser är "godtagbar". Detta eftersom denna sannolikhet i princip inte har med konstruktionens säkerhet att göra, utan egentligen endast påverkar projektets slutkostnad. Om det skulle visa sig att sannolikheten för överskriden larmgräns är större än väntat, gör ju detta endast att de förberedda åtgär-derna måste sättas in med större sannolikhet. Konstruktionens säkerhet ska dock inte kunna påverkas av detta, men det förutsätter förstås att man har organiserat åtgärds-planen så att den är möjlig att skala upp, exempelvis ifall samma typ av åtgärd måste införas på flera ställen samtidigt i ett kritiskt skede.**3.4 Effektiva simuleringsmetoder för att beräkna brottsannolikheter**

3.4.1 Översikt

Så som nämndes ovan är det ofta fullt möjligt att använda det sannolikhetsbaserade ramverket tillsammans med vanlig Monte Carlo-simulering. När man behöver kombi-nera användningen av ramverket med numerisk modellering kan det dock vara lämpligt att använda en effektivare simuleringsmetod för att beräkna brottsannolikheter. Forskningen har inom detta område gått starkt framåt under de senaste decennierna och en rad alternativa metoder finns tillgängliga. Bland dessa kan särskilt nämnas de metoder som på engelska kallas Importance Sampling ("Viktad simulering") och Subset Simulation ("Delmängdssimulering"). Dessa metoder bygger på algoritmer som stegvis söker sig fram mot gränstillståndet, vilket gör att man kan fokusera de utförda simule-ringarna i detta intressanta område av utfallsrummet, till skillnad från vanlig Monte Carlo simulering som ju per definition bara hittar "brott" med en sannolikhet motsva-rande den normalt väldigt låga brottsannolikheten.

I denna rapport har vi använt oss av Subset-simulering. I det följande redogör vi över-siktligt för principen för denna simuleringsmetod. För detaljer kring praktisk implemen-tering av Subset-simulering rekommenderar vi den utmärkta boken av Au & Wang (2014). Färdig Matlab-kod för Subset-simulering finns också fritt tillgänglig på Technische Universität Münchens hemsida (TUM 2022).

3.4.2 Subset-simulering – konceptuell idé

Jämfört med en vanlig Monte Carlo-simulering (se ekvation (5) ovan), så beräknas i en Subset-simulering i stället p_F som en produkt av ett antal större betingade sannolikheter.

Detta låter sig göras genom att definiera p_F med hjälp av ett antal (*M* stycken) nästlade mellanliggande "brotthändelser" F_k med förbestämda brottsannolikheter, p_0. Formellt skrivs detta $F_0 \supset F_1 \supset \cdots \supset F_M$, där F_0 motsvarar en säker händelse (d.v.s. att där inte finns någon sannolikhet alls för brott) och F_M är den faktiska brottgränsen. Brottsannolikheten kan därför omdefinieras som

$$p_F = P\left(\bigcap_{k=1}^{M} F_k\right) = \prod_{k=1}^{M} P(F_k|F_{k-1}), \tag{9}$$

De mellanliggande brotthändelserna definieras genom att sätta $F_k = \{G(\boldsymbol{X}) \leq c_k\}$, där c_k är en konstant större än eller lika med 0. Med hjälp av så kallade Markov-kedjor kan man stegvis, genom att minska c_k, tvinga fram simuleringar allt närmare den faktiska brottgränsen, för vilken $c_k = 0$ per definition. I varje steg används som frön ("seeds") till Markov-kedjorna de simuleringar som i det föregående steget gav brott i den mellanliggande brotthändelsen. Principen för denna stegvisa beräkning illustreras i Figur 3. Eftersom p_0 är en förbestämd sannolikhet, ofta ansatt till 0.1, kan p_F beräknas som

$$p_F \approx \hat{p}_F = p_0^{M-1} p_M, \tag{10}$$

där p_M är en uppskattning av den sista betingade sannolikheten i ekvation (9), vilken ges av

$$p_M = \frac{1}{N_{\text{sub}}} \sum_{i=1}^{N_{\text{sub}}} I_F(\mathbf{x}_{M-1}). \tag{11}$$

Där $\mathbf{x}_{M-1}$ är de utslumpade värdena i det sista steget, som alltså jämförs med den faktiska brottgränsen $G(\boldsymbol{X}) \leq 0$ i indikatorfunktionen I_F, och N_{sub} är antalet simuleringar i varje steg i Subset-simuleringen. Notera likheten med vanlig Monte Carlo-simulering (ekv. (5)).

I Artikel I presenterar vi en algoritm baserad på Subset-simulering, som effektivt tar fram gränser för acceptabelt beteende, $x_{\text{alarm},1}$, genom att iterativt lösa ekvation (6). I Artikel II används algoritmen till ett praktiskt exempel med en sprutbetongförstärkt tunnel.
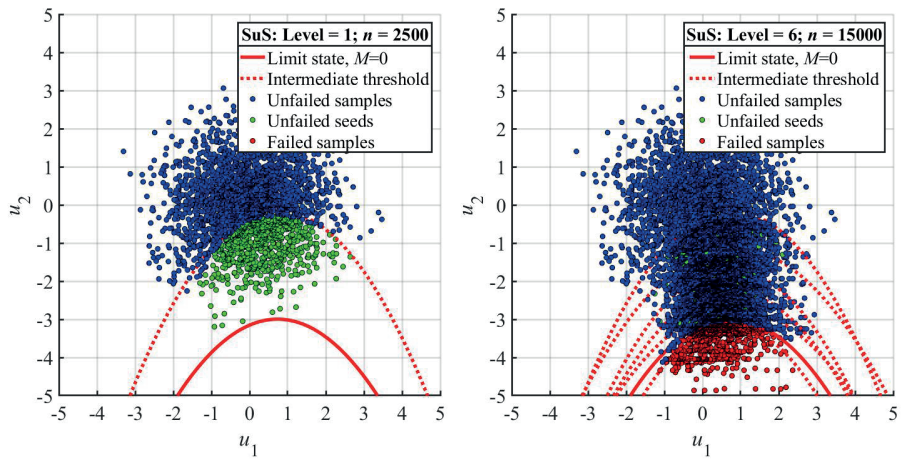
Figur 3. *Subset-simulering.* Vänster*: Åsättande av första mellanliggande gränstillstånd F₁ genom att välja ut de simuleringar (gröna) som ligger närmast det riktiga gränstillståndet, baserat på en första vanlig Monte Carlo-simulering (F₀). Dessa används sedan för att generera nya slumpade utfall som ligger närmare gränstillståndet.* Höger*: Efter 6 upprepningar med mellanliggande gränstillstånd har man lyckats framtvinga simuleringar (röda) som ger brott. (Figurer: Davi R. Damasceno).*

4. REKOMMENDATIONER FÖR PRAKTISK TILLÄMPNING

Dessa rekommendationer baserar sig på våra iakttagelser och erfarenheter från arbetet med artiklarna som bifogats denna forskningsrapport. För att få ut mesta möjliga av nedanstående diskussion rekommenderar vi att man först läser de bifogade artiklarna.

4.1 Vilken definition av observationsmetoden är lämplig?

Att användning av observationsmetodens principer är en central del av dagens berg-byggande brukar ofta framhållas i diskussioner kring dimensioneringsprinciper för tunnelbyggande. Som exempel på detta kan man ta Trafikverkets skrift "Projektering av bergkonstruktioner" (Trafikverket 2019), där observationsmetoden är den dimension-eringsstrategi som rekommenderas. Det ska dock nämnas att observationsmetodens principer kan definieras olika strikt. I Sverige har man traditionellt använt sig av en mindre strikt definition, som gått under namnet Aktiv design (Stille 1986). Aktiv design, med sina enkla men vaga principer (1) förutsägelse, (2) observation och (3) åtgärd, ger användaren ett större mått av frihet och tolkningsutrymme, än den mer strikta definitionen i Eurokoden. Denna frihet lägger samtidigt ett större ansvar på ingenjören, då det inte finns explicit definierade tillvägagångssätt, möjligen annat än det som utarbetats till branschpraxis. Eftersom vi idag ser att Eurokodens användnings-område sannolikt inom kort kommer att innefatta även bergkonstruktioner, anser vi att det är relevant att i denna rapport utgå ifrån Eurokodens definition av observations-metoden snarare än principerna för Aktiv design. Det sannolikhetsbaserade ramverket för observationsmetoden erbjuder ett sätt att uppfylla de krav som Eurokoden ställer.

4.2 Val av simuleringsmetod

För att använda observationsmetoden i det sannolikhetsbaserade ramverket krävs en del statistiska förkunskaper, särskilt om mer avancerade simuleringsmetoder än vanlig Monte Carlo ska användas. Subset-simulering är dock inte mer komplicerat än att det kan läras ut i kurser på masternivå inom anläggningsbyggande. Sådana kurser finns exempelvis på Technische Universität München. Med hjälp av deras fritt tillgängliga online-resurser och den utmärkta boken av Au & Wang (2014) är självstudier en reell möjlighet för den intresserade, om man har i dag har kunskaper om hur vanlig Monte Carlo simulering fungerar.

Erfarenheterna från detta forskningsprojekt visar att vanlig Monte Carlo simulering räcker gott ifall den aktuella gränsfunktionen kan beskrivas analytiskt. Även enkla FE-modeller kan till nöds hanteras med vanlig Monte Carlo, men beräkningstiden kan snabbt stiga, särskilt ifall man ska ned på mycket låga tillåtna brottsannolikheter. För

mer komplicerade FE-modeller, såsom den som det tunneltvärsnitt som utvärderas i Artikel II, kommer beräkningstiderna att bli ohanterliga med vanlig Monte Carlo, i alla fall om man använder en vanlig persondator. För dessa fall är Subset-simulering klart att rekommendera, och även med den metoden kan beräkningarna komma att räknas i dagar.

Det ska dock påpekas att beräkningstiden även styrs av begärd noggrannhet i fastställd larmgräns; snabbare, överslagsmässiga körningar för att undersöka det generella beteendet hos konstruktionen kan enkelt göras genom att välja färre antal simuleringar och genom att öka toleransen för fel i genererad larmgräns. När man valt dimensioner på förstärkning och sett att dessa genererar en rimlig storleksordning på sannolikheten att larmgränsen överskrids (ekv. (8)), kan en mer noggrann beräkning göras för att slutgiltigt fastställa larmgränsens värde.

4.3   Beräkning av brottsannolikhet baserad på resultat från FE-modell

Att beräkna brottsannolikheter kräver ett visst mått av kunskaper i något programmeringsspråk. I detta forskningsprojekt har vi utfört beräkningarna i både Matlab och multifysikprogramvaran COMSOL. I Artikel I utfördes Monte Carlo- och subsetsimuleringarna i Matlab, till vilka beräkningsresultat från en FE-modell i COMSOL importerades när en gränsfunktion behövde utvärderas. I Artikel II implementerades algoritmen för framtagande av larmgränser fullständigt i COMSOL. Detta krävde dock färdigheter i Java-programmering. Genom vidare utvecklingsarbete bör det dock vara möjligt att generalisera algoritmen för framtagande av larmgränser fullständigt och därmed integrera algoritmen helt och hållet i programvaror såsom COMSOL. Användaren skulle därmed inte längre behöva koda algoritmen själv. Detta skulle klart öka användarvänligheten av det sannolikhetsbaserade ramverket i stort och algoritmen för larmgränser mer specifikt. (Jämför med skillnaden i användarvänlighet av att använda ett FEM-program, med att koda FEM-beräkningen själv).

4.4   Implementering inom tunnelbyggnad

I Artikel II visar vi hur sannolikhetsbaserade larmgränser kan tas fram för en sprutbetonglinings deformation vid tunneldrivning i svåra geotekniska förhållanden. Vi diskuterar även hur dessa larmgränser kan användas i observationsmetodens ramverk. En utmaning inom just tunnelbyggnad är svårigheten att förutsäga hur bergmassan avlastas allteftersom tunnelfronten drivs framåt. Därmed blir det stor osäkerhet kring vilken belastning som förstärkningen kommer att behöva klara efter installation. Denna osäkerhet beror dels på den bristande kännedomen om de geotekniska förhållandena, dels på svårigheten att korrekt modellera själva avlastningen. Osäkerheten i geotekniska

förhållanden kan bara minskas genom bättre undersökningar av de geologiska förhållandena. Avseende modelleringen måste man här göra ett avvägande mellan mer exakta, men mycket beräkningstunga 3D-modeller och förenklade men snabbare 2D-modeller. I Artikel II föreslår vi ett praktiskt tillvägagångssätt för att skatta osäkerheten som 2D-modellen medför avseende vid vilken grad av avlastning som förstärkningen installeras. Vårt tillvägagångssätt ska ses som en ingenjörsmässig lösning på ett svårt beräkningstekniskt problem. När noggrannheten i detta förfarande inte bedöms tillräckligt, rekommenderas fördjupade analyser med 3D-modeller, genom användande av så kallade Longitudinella deformationsprofiler (se Vlachopoulos & Diederichs 2009).

En annan beräkningsteknisk utmaning inträffar ifall man önskar fastställa mer än en larmgräns för att övervaka ett brottgränstillstånd. Man kan exempelvis tänka sig att man vill ha en larmgräns för tunneltakets deformation och en larmgräns för tunnelväggens deformation. Ekvationen för framtagandet av dessa larmgränser blir då underbestämd, såsom noterades i avsnitt 3.3.4. Det innebär att man får många möjliga kombinationer av larmgränser som uppfyller ekvationen. Den optimala kombinationen blir då ett beslutsteoretiskt optimeringsproblem: Ska man tillåta mycket deformation i taket och lite i väggen, eller tvärtom? Eller är det bättre att tillåta medelstor deformation i båda? Med hjälp av ekvation (8) kan man beräkna sannolikheten för att respektive föreslagen larmgräns överskrids. Därmed kan man bedöma ifall respektive föreslagen larmgräns är rimlig med avseende på förväntade kostnader för åtgärder som behöver sättas in, ifall larmgränsen överskrids. Ur ett beslutsteoretiskt perspektiv finns det där potential att optimera val av larmgräns med avseende på vilka kostnader som överskridande av larmgränsen skulle medföra. Alltså: ifall extra förstärkningsåtgärder i taket är dyrare än motsvarande åtgärd i väggen, kan det finnas skäl att göra larmgränsen för taket mer tillåtande och kompensera detta genom att göra larmgränsen för väggen mer strikt. Formellt kan denna optimering på ett generellt plan beskrivas med ekvation (3).

4.5 När är det sannolikhetsbaserade ramverket lämpligt att använda?

Såsom framgått i denna rapport medför användandet av det sannolikhetsbaserade ramverket för framtagande av larmgränser en relativt lång startsträcka, åtminstone avseende kompetensutveckling och programmering av algoritmen vid första användningstillfället. När man väl har utfört en beräkning av en larmgräns ser man dock snabbt att beräkningarna är tämligen standardiserade. Den algoritm som presenteras i Artikel I är generell, så till vida att så länge man kan beskriva det aktuella gränstillståndet med en gränsfunktion, så kan en larmgräns tas fram. Den ingenjörsmässiga utmaningen ligger då främst i att korrekt sätta upp FE-modellen, definiera ingående geotekniska parametrar som sannolikhetsfunktioner och korrekt beakta mät- och transformationsfel, samt ta fram lämpligt koncept för hur observationerna ska utföras. En del av dessa utmaningar måste förstås lösas även ifall man inte använder sannolikhetsbaserade metoder.

Vår bedömning är att det extra steg som krävs för att utföra de sannolikhetsbaserade analyserna är rimligt att ta främst vid kritiska dimensioneringssituationer, det vill säga när både osäkerheterna och konsekvenserna vid brott är stora. Sannolikhetsbaserade beräkningsmetoder erbjuder då en god möjlighet att på djupet analysera storleken på de risker som man tar med olika designlösningar. Inte minst erbjuder det sannolikhets-baserade ramverket en möjlighet att bedöma sannolikheten att en vald larmgräns kommer att överskridas under byggtiden. Sådan information kan vara av stor betydelse vid val av designlösning, eftersom en stor sannolikhet för överskriden larmgräns indike-rar ett stort ekonomiskt risktagande kopplat till kostsamma åtgärder, som behöver sättas in för att undvika brott i konstruktionen.

5.   SLUTORD

I detta forskningsprojekt har vi undersökt hur man kan kombinera observationsmetoden med sannolikhetsbaserade dimensioneringsprinciper. Utgångspunkten har legat i att bättre tillgängliggöra det sannolikhetsbaserade ramverk som togs fram i Spross (2016) doktorandprojekt, samt visa på dess praktiska tillämpbarhet. Eftersom observations-metodens principer idag är en central del av dimensioneringsarbetet inom berg-byggande, bör det finnas ett stort behov av praktiska verktyg som hjälper till i detta arbete. Detta eftersom observationsmetoden i sig är definierad på ett generellt plan i Eurookod 7, vilket lämnar ett stort ansvar på den ansvariga ingenjören. Dessutom är till-vägagångssättet för uppfyllandet av samhällets säkerhetskrav inte tydligt angett i Eurokodens avsnitt om observationsmetoden.

I forskningsprojektet har vi främst fokuserat på hur man tar fram gränser för acceptabelt beteende ("larmgränser"), som ska avgöra när (förstärknings-)åtgärder behöver sättas in för att undvika brott i konstruktionen. En generell beräkningsalgoritm för framtagandet av sådana larmgränser utvecklades i Artikel I. Användbarheten av algoritmen visades i ett större praktiskt exempel i Artikel II. I exemplet visades hur man kan fastställa larm-gränser för acceptabel deformation i en sprutbetonglining. Som utgångspunkt för exemplet användes de geotekniska förhållandena som observerats vid de geotekniska undersökningarna inför drivningen av Förbifart Stockholms tunnlar under Mälaren, om än med några förenklingar av praktiska skäl.

Baserat på våra erfarenheter från forskningsprojektet, anser vi att det sannolikhets-baserade ramverket för observationsmetoden (och algoritmen för larmgränser i synner-het) kan utgöra ett bra tillskott till verktygslådan för den som dimensionerar berganläggg-ningar i svåra geotekniska förhållanden, samt när man ställs inför stora risker. Observat-ionsmetoden kan, på grund av sin generellt formulerade definition, annars lätt miss-brukas genom otillräcklig kompetens eller otillbörligt risktagande. Att i sådana fall istället hantera geotekniska osäkerheter på ett stringent sätt i ett sannolikhetsbaserat ramverk är ett sätt att öka tillförlitligheten i de föreslagna designlösningarna och tydlig-göra för ansvariga beslutsfattare vilka risker man står inför. Vi vill hävda att man genom att ge god grund för välinformerade beslut uppnår ett mer kostnadseffektivt byggande.

6. REFERENSER

Au, S.-K. & Wang, Y. 2014. *Engineering risk assessment with subset simulation.* Singapore: John Wiley & Sons.

Baecher, G.B. & Christian, J.T. 2003. *Reliability and statistics in geotechnical engineering.* Chichester: John Wiley & Sons.

Bayes, T. 1763. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F.R.S. communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S. *Philosophical Transactions, 53*, 370-418.

Benjamin, J.R. & Cornell, C.A. 1970. *Probability, statistics, and decision for civil engineers.* New York, NY: McGraw–Hill Book Company.

Bjureland, W., Spross, J., Johansson, F., Prästings, A. & Larsson, S. 2017. Reliability aspects of rock tunnel design with the observational method. *International Journal of Rock Mechanics and Mining Sciences, 98*, 102-110.

CEN 2004. *EN 1997-1:2004 Eurocode 7: Geotechnical design – Part 1: General rules.* Brussels: European Committee for Standardisation.

Frank, R., Bauduin, C., Driscoll, R., Kavvadas, M., Krebs Ovesen, N., Orr, T. & Schuppener, B. 2004. *Designers' guide to EN 1997-1 Eurocode 7: Geotechnical design-General rules.* London: Thomas Telford.

Gothäll, R. 2011. *Analys av konvergensmätningar. Rapport 109.* Stockholm: BeFo.

Herodotus [translated by A. D. Godley 1922] c. 430 B.C. Books V-VII. In: T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post & E. H. Warmington (eds.), *The histories.* London: William Heinemann.

Holmberg, M. & Stille, H. 2007. *Observationsmetodens grunder och dess tillämpning på design av konstruktioner i berg.* Stockholm: BeFo.

Holmberg, M. & Stille, H. 2009. *Observationsmetoden och deformationsmätningar vid tunnelbyggande [The observational method and deformation measurements in tunnels]. Report 93.* Stockholm: BeFo.

Kadefors, A. & Bröchner, J. 2008. *The observational method for underground excavations: contracts and collaboration.* Stockholm: BeFo Rock Engineering Research Foundation.

Maidl, U., Köhler, M. & Schretter, K. 2011. Implementation of the observational method in mechanised tunnelling–contracts H3-4 and H8 in the Lower Inn Valley. *Geomechanics and Tunnelling, 4*(5), 405-413.

Miranda, T., Dias, D., Pinheiro, M. & Eclaircy-Caudron, S. 2015. Methodology for real-time adaption of tunnels support using the observational method. *Geomechanics and Engineering, 8*(2), 153-171.

Müller, R., Larsson, S. & Spross, J. 2014. Extended multivariate approach for uncertainty reduction in the assessment of undrained shear strength in clays. *Canadian Geotechnical Journal, 51*, 231-245.

Peck, R.B. 1969. Advantages and limitations of the observational method in applied soil mechanics. *Géotechnique, 19*(2), 171-187.

Phoon, K.-K. & Kulhawy, F.H. 1999a. Characterization of geotechnical variability. *Canadian Geotechnical Journal, 36*(4), 612-624.

Phoon, K.-K. & Kulhawy, F.H. 1999b. Evaluation of geotechnical property variability. *Canadian Geotechnical Journal, 36*(4), 625-639.

Rackwitz, R. & Schrupp, K. 1985. Quality control, proof testing and structural reliability. *Structural Safety, 2*(3), 239-244.

Raiffa, H. & Schlaifer, R. 1961. *Applied statistical decision theory.* Boston, MA: Division of Research, Harvard Business School.

Spross, J. 2016. *Toward a reliability framework for the observational method.* Ph.D. thesis, TRITA-JOB 1023. Stockholm: KTH Royal Institute of Technology.

Spross, J. 2017. Vilken roll har observationsmetoden i framtidens bergbyggande? In: E. Friedman (ed.), *Artikel till presentation på Bergmekanikdagen, Stockholm, 13 mars 2017*. Stockholm: BeFo.

Spross, J. & Johansson, F. 2017. When is the observational method in geotechnical engineering favourable? *Structural Safety, 66*, 17-26.

Spross, J., Johansson, F., Uotinen, L.K.T. & Rafi, J.Y. 2016. Using observational method to manage safety aspects of remedial grouting of concrete dam foundations. *Geotechnical and Geological Engineering, 34*, 1613-1630.

Spross, J. & Larsson, S. 2014. On the observational method for groundwater control in the Northern Link tunnel project, Stockholm, Sweden. *Bulletin of Engineering Geology and the Environment 73*(2), 401-408.

Spross, J., Stille, H., Johansson, F. & Palmstrøm, A. 2018. On the need for a risk-based framework in Eurocode 7 to facilitate design of underground openings in rock. *Rock Mechanics and Rock Engineering 51*(8), 2427–2431.

Spross, J., Stille, H., Johansson, F. & Palmstrøm, A. 2019. *Utredning av riskbaserade principer inom bergdimensionering: Så bör en standard vara uppbyggd. Report 190.* Stockholm: BeFo.

Spross, J., Stille, H., Johansson, F. & Palmstrøm, A. 2020. Principles of risk-based rock engineering design. *Rock Mechanics and Rock Engineering, 53*, 1129–1143.

Statens Järnvägar. 1922. *Statens Järnvägars geotekniska kommission 1914-22: Slutbetänkande avgivet till Kungl. Järnvägsstyrelsen den 31 maj 1922 [The Geotechnical Committee of the State Railways 1914-1922: Final report presented to the Royal Board of State Railways on 31 May 1922].* Stockholm: Statens Järnvägar.

Stille, H. 1986. Experiences of design of large caverns in Sweden. In: K. H. O. Saari (ed.), *Proceedings of the International Symposium on Large Rock Caverns, Helsinki, 25-28 August 1986*. Oxford, UK: Pergamon Press, 231-241.

Stille, H. & Holmberg, M. 2010. Examples of applications of observational method in tunnelling. *Geomechanics and Tunnelling, 3*(1), 77-82.

Trafikverket. 2019. *Projektering av bergskonstruktioner, nr 2019:062.* Sundbyberg: Trafikverket.

TUM. 2022. *Subset Simulation for reliability analysis* [Online]. München: TUM. Available: https://www.cee.ed.tum.de/en/era/software/reliability/subset-simulation/ [Accessed 2022-01-12.

van der Krogt, M.G., Schweckendiek, T. & Kok, M. 2019. Uncertainty in spatial average undrained shear strength with a site-specific transformation model. *Georisk, 13*(3), 226-236.

Vlachopoulos, N. & Diederichs, M.S. 2009. Improved longitudinal displacement profiles for convergence confinement analysis of deep tunnels. *Rock Mechanics and Rock Engineering 42*(2), 131-146.

von Neumann, J. & Morgenstern, O. 1944. *Theory of games and economic behaviour.* Princeton: Princeton University Press.

Vrouwenvelder, A.C.W.M. 2002. Developments towards full probabilistic design codes. *Structural Safety, 24*(2), 417-432.

Zetterlund, M. 2014. *Value of information analysis in rock engineering investigations.* Ph.D. thesis. Gothenburg: Chalmers University of Technology.

7. BILAGOR

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

Johan Spross[a,*], Tobias Gasch[b]

[a] *Division of Soil and Rock Mechanics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden*
[b] *Division of Concrete Structures, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden*

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 (J. Spross), tobias.gasch@byv.kth.se (T. Gasch).

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
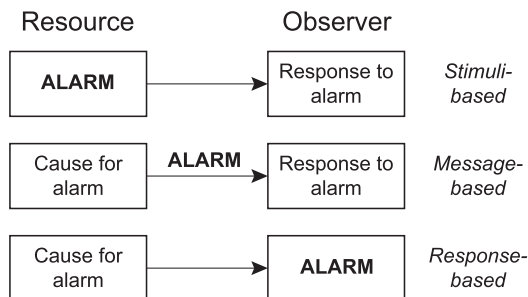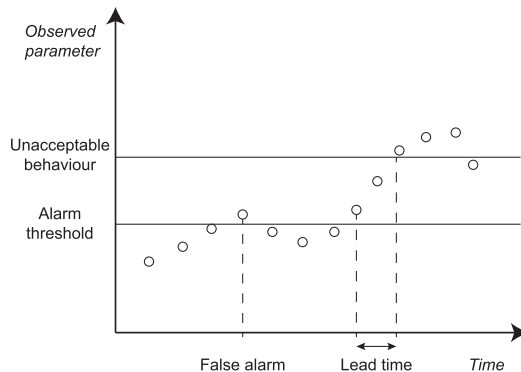

Fig. 2. False alarm and lead-time.

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

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.)

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, $G(\boldsymbol{X})$, where the vector $\boldsymbol{X}$ 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 $F = \{G(\boldsymbol{X}) \leq 0\}$. 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

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

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

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

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

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

where the updated failure region Ω_Z accounts for any reformulations of the limit state functions that the information called for, and $f_{X|Z}$ 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

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

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

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

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

where the function $h(\boldsymbol{X})$ 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 $h(\boldsymbol{X})$ 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

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

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, $x_{1,\text{alarm}}$, provides the information that $h(\boldsymbol{X}) = x_{1,\text{alarm}} - X_1 \leq 0$. (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 $h(\boldsymbol{X}) = X_2 - x_{2,\text{alarm}} \leq 0$, with reference to Fig. 3.)

Using Eq. (7), a value for $x_{1,\text{alarm}}$ 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:

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

This procedure is suggested by Spross and Johansson [26] for
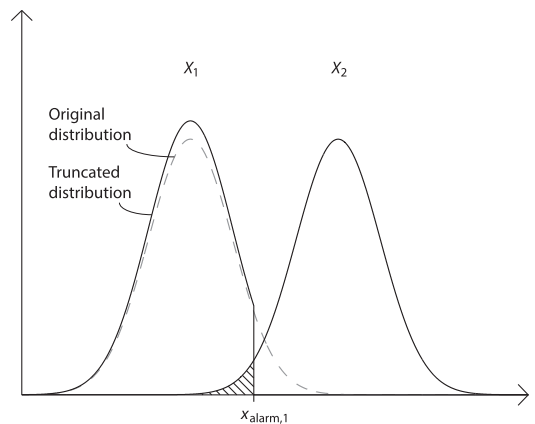
Fig. 3. An alarm threshold against too high readings of X_1 is provided by a truncation of X_1 at its upper end. This implies that the overlapping of the two distributions is significantly reduced and so is the corresponding failure probability. Correspondingly, an alarm threshold against too low readings 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 $x_{1,\text{alarm}}$ 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

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

where the vector $\boldsymbol{x}_{\text{alarm}}$ 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 $\boldsymbol{x}_{\text{alarm}}$, 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 ε is used:

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

where ∇ is the gradient operator, and ∇^{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 $\mathbf{f}_{\text{b}}$ acting over an infinitesimal volume must be balanced by the change in stress given by the second-order symmetric tensor σ. This must strictly hold for every point in domain Λ; i.e., we have that

$$\mathbf{0} = \nabla{\cdot}\boldsymbol{\sigma} + \mathbf{f}_{\text{b}} \text{ in } \Lambda. \tag{11}$$

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:

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

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, ν. 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 Γ_1 by directly prescribing the displacements to be $\mathbf{u}_0$:

$$\mathbf{u} = \mathbf{u}_0 \text{ on } \Gamma_1. \tag{13}$$

Alternatively, Neumann type boundary conditions can be used by prescribing the surface traction $\mathbf{t}$ on surface Γ_2, such that

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

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 Λ. This conflicts with the approxamitive 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 Λ. Then, using integration-by-parts and the divergence theorem to remove second order derivatives, we arrive at

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

which now holds for all test functions, instead of as previously for every point in Λ, 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

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

whereas vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ 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:

$$\mathbf{0} = -\hat{\mathbf{v}}^{\text{T}} \int_{\Lambda} (\mathbf{LN})^{\text{T}}\mathbf{C}(E, \nu)\mathbf{LN}dV\hat{\mathbf{u}} + \hat{\mathbf{v}}^{\text{T}} \int_{\Gamma} \mathbf{N}^{\text{T}}\mathbf{t}dS + \hat{\mathbf{v}}^{\text{T}} \int_{\Lambda} \mathbf{N}^{\text{T}}\mathbf{f}_{\text{b}}dV, \tag{17}$$

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

$$\mathbf{K}\hat{\mathbf{u}} = \mathbf{f}, \tag{18}$$

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 $F_0 \supset F_1 \supset \cdots \supset F_M$, where F_0 is a certain event and $F_M = F$ is the event of interest. Then, we have that

$$p_{\text{F}} = P\left(\bigcap_{k=1}^{M} F_k\right) = \prod_{k=1}^{M} P(F_k|F_{k-1}). \tag{19}$$

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 [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

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

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

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

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 $F_0^* \supset F_1^* \supset \cdots \supset F_M^*$, where $F_k^* = F_k \cap Z$ and $F_0^* = F_0 \cap Z = Z$. Then, we can modify Eq. (19) into

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

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, τ, 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, τ, and κ. (κ 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 κ 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 $\tilde{\mathbf{b}}_i = [\tilde{b}_i^{(1)}, \tilde{b}_i^{(l)}, \cdots, \tilde{b}_i^{(N_s)}]$ from a proposal PDF $q_i^*(b_i^{(\text{seed})})$ for each basic parameter in $\mathbf{B}$.
         – Calculate for each proposal $r_i^{(l)} = \frac{q_i^*(\tilde{b}_i^{(l)};b_i^{(\text{seed})})f_{B_i}(\tilde{b}_i^{(l)})}{q_i^*(b_i^{(\text{seed})};\tilde{b}_i^{(l)})f_{B_i}(b_i^{(\text{seed})})}$.
         – Generate $u_i^{(l)}$ uniformly distributed on $[0,1]$ for each proposal.
         – Set for all l and all basic parameters, $\tilde{\tilde{b}}_i^{(l)} = \begin{cases} \tilde{b}_i^{(l)} & \text{if } u_i^{(l)} < r_i^{(l)} \\ b_i^{(\text{seed})} & \text{otherwise} \end{cases}$ and collect them in $\tilde{\tilde{\mathbf{b}}}_k$.
      ii. Run the structural model for each set of proposed samples in $\tilde{\tilde{\mathbf{b}}}_k$ to find the complete set of proposed samples, $\tilde{\tilde{\mathbf{x}}}_k$.
      iii. Set, for all N parameter sets of proposed samples, $\mathbf{x}_k = \begin{cases} \tilde{\tilde{\mathbf{x}}}_k & \text{if } \tilde{\tilde{\mathbf{x}}}_k \in F_k^* \\ \mathbf{x}_k^{(\text{seed})} & \text{otherwise} \end{cases}$
      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 κ 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 κ will depend on the number of rejected sets of samples in the truncation. The choice of τ 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, $x_{1,\text{alarm}}^{(1)}$, is made. In principle, $x_{1,\text{alarm}}^{(1)}$ can be set anywhere in the generated sample $\boldsymbol{x}_1$. For load-related parameters, we suggest to set $x_{1,\text{alarm}}^{(1)}$ in the range $[\bar{\boldsymbol{x}}_1, \max \boldsymbol{x}_1]$, where $\bar{\boldsymbol{x}}_1$ is the sample mean, and for resistance-related parameters, we suggest the range $[\min \boldsymbol{x}_1, \bar{\boldsymbol{x}}_1]$. Note that if the alarm threshold is needed in the other end of the range of $\boldsymbol{x}_1$ than suggested above, in order to satisfy $p_{\text{F,T}}$, the probability of violating $x_{1,\text{alarm}}$, i.e. $P(X_1 > x_{1,\text{alarm}})$ 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 $x_{1,\text{alarm}}^{(j)}$. As the calculated $\hat{p}_{\text{F}}$ indicates whether $x_{1,\text{alarm}}^{(j)}$ is set too high or too low, each iteration will reduce the possible range of the final $x_{1,\text{alarm}}$. 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 $x_{1,\text{alarm}}^{(j+1)}$ 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 $\widetilde{\boldsymbol{b}}_i$ 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 $\widetilde{\boldsymbol{b}}_i$ be normally distributed with mean $\widetilde{\mu}_{\text{b},i} = b_i^{(\text{seed})}$ and standard deviation $\widetilde{\sigma}_{\text{b},i} = \sigma_{\text{b},i}$, 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^2	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 $\boldsymbol{B}$ 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 ν 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

$$G(\boldsymbol{X}) = \frac{f_\text{t}}{E} - \varepsilon \le 0, \tag{23}$$

where ε is the maximum strain in the beam subjected to q. For a simply supported beam, ε is always located at the bottom material fibre in the mid-section. The sought alarm threshold in terms of strain is denoted $\varepsilon_{\text{alarm}}$, 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 ε, we define from structural mechanics the maximum field moment for a simply supported beam as

$$M_\text{f} = \frac{ql^2}{8d}, \tag{24}$$

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

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

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
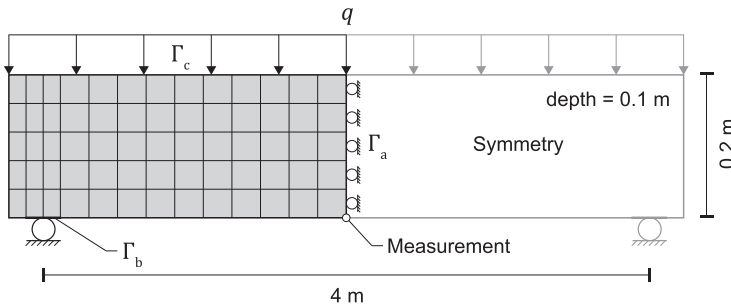

Fig. 4. Schematic description of the analysed concrete beam including dimensions, finite element discretization, and boundary conditions.

$$u = \frac{5ql^4}{384EId}. \tag{26}$$

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 Γ_a, where the displacement component normal to Γ_a is constrained to zero. The roller boundary condition of surface Γ_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 Γ_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 Γ_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 0.04×0.04 m. Each finite element is in the example defined using quadratic serendipity shape functions and thus $\mathbf{N}(\mathbb{X})$ becomes a 16×16 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 ε_{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 τ 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

Table 3
Simulation constants used in the example.

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

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_{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_{alarm} is affected mainly by the number of Subset simulations, N. Increasing N reduces the variability in u_{alarm}; for example, increasing N from 5000 to 10 000 implies a reduction in coefficient of variation for u_{alarm} from 4.7% to 3.1% (based on 50 calculations of u_{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
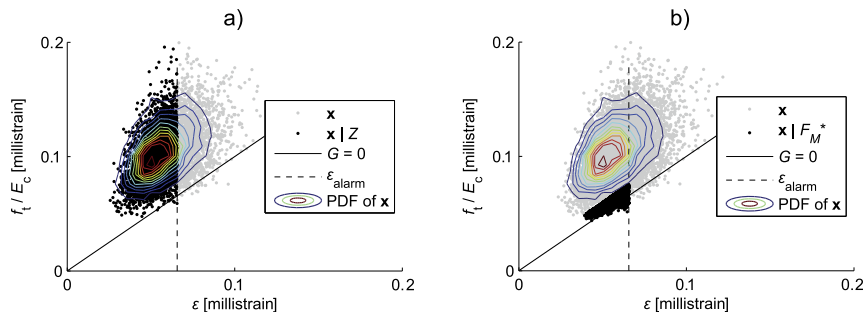
Fig. 5. Simulation results for the concrete beam. a) The remaining sample (in black), given that $\varepsilon < \varepsilon_{\text{alarm}}$. b) The last generated subset sample, given that $G(X) < c_k$ and $\varepsilon < \varepsilon_{\text{alarm}}$.

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
$\varepsilon_{\text{alarm}}$ [millistr]	0.0670	0.0678	0.0661	0.0659
u_{alarm} [mm]	1.11	1.11	1.10	1.10
t_{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 $p_{\text{F,T}}$ 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: $G(X, t) = 0$. 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 Δt_{lead}, if contingency actions were put into operation at t_i), Eq. (8) can be extended to

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

Having established $x_{1,\text{alarm}}(t_i + \Delta t_{\text{lead}})$ with the proposed algorithm (Table 1) as a percentile P_{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 $p_{\text{F,T}}$.

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 $Z = \{h(X) = 0\}$, instead of treating the data as inequality information (Eq. (5)), the updated failure probability $p_{\text{F|Z}}$ would potentially be more reduced. We believe, however, that the use of alarm thresholds has a practical advantage over direct computation of $p_{\text{F|Z}}$, 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.
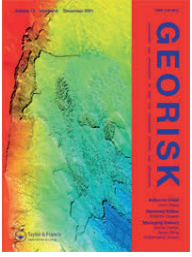
Acknowledgements

References

[1] Brownjohn JMW. Structural health monitoring of civil infrastructure. Philos Trans R Soc A 2007;365:589–622.

[2] Paté-Cornell ME, Tagaras G. Risk costs for new dams: economic analysis and effects of monitoring. Water Resour Res 1986;22:5–14.

[3] De Sortis A, Paoliani P. Statistical analysis and structural identification in concrete dam monitoring. Eng Struct 2007;29:110–20.

[4] Glišic B, Inaudi D, Vurpillot S. Structural monitoring of concrete structures. Como, Italy: Wiley; 2002. p. 1–10.

[5] Buckby RJ, Key CL, Palmer MJ, Penman JG, Swannell NG. An observational approach to design of a new grout curtain at Wimbleball Dam. In: Winter MG, Smith DM, Eldred PJL, Toll DG, editors. Proceedings of the XVI ECSMGE: geotechnical engineering for infrastructure and development. London: ICE Publishing; 2015. p. 559–654.

[6] 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.

[7] Smith LM. In-service monitoring of nuclear-safety-related structures. Struct Engineer 1996;74:210–1.

[8] 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.

[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.

[10] 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.

[11] Müller R, Larsson S, Spross J. Multivariate stability assessment during staged construction. Can Geotech J 2016;53:603–18.

[12] Schweckendiek T, Vrouwenvelder ACWM, Calle EOF. Updating piping reliability with field performance observations. Struct Saf 2014;47:13–23.

[13] 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.

[14] 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.

[15] 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.

[16] Luque J, Straub D. Reliability analysis and updating of deteriorating systems with dynamic Bayesian networks. Struct Saf 2016;62:34–46.

[17] 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.

[18] CEN. EN 1997-1:2004. Eurocode 7: geotechnical design – Part 1: general rules. Brussels: European Committee for Standardisation; 2004.

[19] Peck RB. Advantages and limitations of the observational method in applied soil mechanics. Géotechnique 1969;19:171–87.

[20] Huang C, Shachter R. Alarms for monitoring: a decision-theoretic framework. Technical Report SMI-97-0664. Stanford: Section on Medical Informatics, Stanford University School of Medicine; 1997.

[21] Frank R, Bauduin C, Driscoll R, Kavvadas M, Krebs Ovesen N, Orr T, et al. Designers' guide to EN 1997–1 Eurocode 7: geotechnical design – general rules. London: Thomas Telford; 2004.

[22] Powderham AJ, Nicholson DP. The way forward. In: Nicholson DP, editor. The observational method in geotechnical engineering. London: Thomas Telford; 1996. p. 195–204.

[23] Powderham AJ. The observational method—learning from projects. Proc Inst Civ Eng Geotech 2002;155:59–69.

[24] Spross J, Johansson F, Stille H, Larsson S. Towards an improved observational method. In: Alejano LR, Perucho A, Olalla C, Jiménez R, editors. EUROCK 2014. London: Taylor & Francis; 2014. p. 1435–40.

[25] 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.

[26] Spross J, Johansson F. When is the observational method in geotechnical engineering favourable? Struct Saf 2017;66:17–26.

[27] Wallin S. Chasing a definition of "alarm". J Netw Syst Manage 2009;17:457–81.

[28] Breznitz S. Cry wolf: the psychology of false alarms. Hillsdale: Lawrence Erlbaum Associates; 1984.

[29] 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.

[30] Sättele M, Krautblatter M, Bründl M, Straub D. Forecasting rock slope failure: how reliable and effective are warning systems? Landslides 2016;13:737–50.

[31] Paté-Cornell ME, Benito-Claudio CP. Warning systems: response models and optimization. In: Covello 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.

[32] Olsson L, Stille H. Observation systems with alarm thresholds and their use in designing underground facilities. Stockholm: SKB; 2002.

[33] Der Kiureghian A, Ditlevsen O. Aleatory or epistemic? Does it matter? Struct Saf 2009;31:105–12.

[34] 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.

[35] Prästings A, Müller R, Larsson S. The observational method applied to a high embankment founded on sulphide clay. Eng Geol 2014;181:112–23.

[36] Wu TH. 2008 Peck lecture: the observational method: case history and models. J Geotech Geoenviron Eng 2011;137:862–73.

[37] Vrouwenvelder ACWM. Developments towards full probabilistic design codes. Struct Saf 2002;24:417–32.

[38] Straub D. Reliability updating with equality information. Prob Eng Mech 2011;26:254–8.

[39] Straub D. Value of information analysis with structural reliability methods. Struct Saf 2014;49:75–85.

[40] Ditlevsen O, Madsen HO. Structural reliability methods. In: Kgs. Lyngbyeditor. Coastal, maritime and structural engineering. Department of Mechanical Engineering, Technical University of Denmark; 2007.

[41] Hall WB. Reliability of service-proven structures. J Struct Eng 1988;114:608–24.

[42] Au S-K, Beck JL. Estimation of small failure probabilities in high dimensions by subset simulation. Prob Eng Mech 2001;16:263–77.

[43] Koutsourelakis PS, Pradlwarter HJ, Schuëller GI. Reliability of structures in high dimensions, part I: algorithms and applications. Prob Eng Mech 2004;19:409–17.

[44] Bucher C. Asymptotic sampling for high-dimensional reliability analysis. Prob Eng Mech 2009;24:504–10.

[45] Zienkiewicz OC, Taylor RL, Zhu JZ. The finite element method: its basis and fundamentals. 7th ed. Oxford: Butterworth-Heinemann; 2013.

[46] Straub D, Papaioannou I, Betz W. Bayesian analysis of rare events. J Comput Phys 2016;314:538–56.

[47] Au S-K, Wang Y. Engineering risk assessment with subset simulation. Singapore: John Wiley & Sons; 2014.

[48] Papaioannou I, Betz W, Zwirglmaier K, Straub D. MCMC algorithms for subset simulation. Prob Eng Mech 2015;41:89–103.

[49] Schneider R, Thöns S, Straub D. Reliability analysis and updating of deteriorating systems with subset simulation. Struct Saf 2017;64:20–36.

[50] 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.

[51] 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.

Implementation of reliability-based thresholds to excavation of shotcrete-supported rock tunnels

Johan Spross, Tobias Gasch & Fredrik Johansson

Submit your article to this journal ↗

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Implementation of reliability-based thresholds to excavation of shotcrete-supported rock tunnels

Johan Spross ●[a], Tobias Gasch ●[b] and Fredrik Johansson ●[a]

[a]Division of Soil and Rock Mechanics, KTH Royal Institute of Technology, Stockholm, Sweden; [b]COMSOL, Stockholm, Sweden

ABSTRACT

Modern tunnelling in rock relies heavily on information from monitoring and observations during construction as a means to reduce the considerable uncertainty that originates from a lack of knowledge about the ground conditions. By formally integrating the monitoring information into the structural safety evaluation, a comprehensive risk-based design framework can be achieved. The establishment of relevant thresholds against unacceptable structural behaviour is a key aspect of this work. This paper presents an extensive application example based on real-case data, showing how a reliability-based threshold can be established to ensure the serviceability of a shotcrete lining in a rock tunnel in a challenging geological setting. The authors address and discuss practical approaches to consider the model uncertainty related to the confinement loss caused by the tunnel front advance, as well as the transformation uncertainty related to the use of indirect monitoring information. The established threshold is discussed in the context of the observational method.

1. Introduction

Monitoring and observation of the structural behaviour during construction are key components of modern tunnelling, because of the considerable prevailing uncertainties of the geological and geotechnical conditions in the ground; this has paved the way for observation-based tunnelling approaches (Schubert 2008; Stille and Holmberg 2010; Spross and Larsson 2014; Palmström and Stille 2015; Bjureland et al. 2017; Spross et al. 2018). These uncertainties originate either from natural variability in the ground conditions (aleatory uncertainty) or from a lack of knowledge of the conditions (epistemic uncertainty), where the latter is caused mainly by the limited extent of the pre-investigations performed. The knowledge gained from monitoring and observations during construction are therefore vital in satisfying the structural safety requirements, as additional information generally implies that the epistemic uncertainties are reduced and that the calculated structural reliability is improved (Spross et al. 2020).

If the tunnel has been designed with reliability-based methods (e.g. Langford and Diederichs 2013; Lü et al. 2017; Kroetz et al. 2018; Napa-García, Beck, and Celestino 2018; Bjureland et al. 2019), the knowledge gained

can be accounted for straightforwardly through Bayesian updating, as demonstrated recently for tunnels by Feng and Jimenez (2015) and Feng et al. (2019), and for other geotechnical structures by Schweckendiek and Vrouwenvelder (2013), Li et al. (2016) and Contreras and Brown (2019). However, the primary purpose of performing monitoring and observations of the structural performance is normally to gain information about when specific measures are needed to ensure that the considered limit states are not attained. This requires that the monitoring or observations be linked to a threshold that triggers the implementation of safety-enhancing modifications to the structure. Consequently, a threshold needs to be established so that the triggered measure can be implemented before the corresponding limit state is attained. From a structural reliability point of view, the threshold shall be established such that the probability of failure of the structural component is satisfactory (i.e. the calculated probability of failure, p_F, is less than the target, $p_{F,T}$), as long as the monitoring result does not violate the threshold.

Although such thresholds often have a key role in the execution of a geotechnical design, there is, however, a considerable lack of practical guidance on how to establish the threshold so that it provides a sufficient safety

margin with respect to the problem at hand; for example, the application guidelines to Eurocode 7 (Frank et al. 2004) state only that "it is the designer's responsibility to prepare and communicate specifications for any such monitoring".

Spross and Johansson (2017) showed in a simplified example how such reliability-based thresholds can be determined and applied to monitor the deformation of a rock pillar in the context of the observational method. The procedure for determining reliability-based thresholds was later theoretically extended by Spross and Gasch (2019) to a more general class of engineering problems, by making use of the Finite Element Method and Subset simulation. There is, however, a need for more comprehensive case examples that discuss the practical implementation of such thresholds to complex geotechnical structures, where the epistemic uncertainty typically is considerable. In this paper, we, therefore, investigate how reliability-based thresholds can be established to ensure the structural reliability of the shotcrete lining in a rock tunnel. In particular, we discuss the practical management of the threshold in the design of tunnels with the observational method (Peck 1969; CEN 2004) and the practical challenges in accounting for the involved model errors. The study uses the challenging geological conditions that were encountered in the Stockholm bypass rock tunnel project (European highway E4), where the excavation passed through a fault zone under lake Mälaren, west of Stockholm, Sweden.

2. Method to establish reliability-based thresholds

2.1. Threshold definition

Describing a limit state as a function $G(\mathbf{X}) = 0$, with $\mathbf{X} = [X_1, \ldots, X_m]$ being a vector of the relevant random variables, Spross and Johansson (2017) showed how a set of threshold values, $\mathbf{x}_{\text{alarm}}$, can be determined from an equality, so that the thresholds facilitate the structural safety requirements:

$$P(G(\mathbf{X}) \leq 0 | \mathbf{X} \leq \mathbf{x}_{\text{alarm}}) = p_{\text{F,T}}. \quad (1)$$

In short, the procedure utilises the fact that a potential monitoring result can be described in terms of a probability of violating the threshold. For example, assuming monitoring of X_1, the probability becomes $P(h(\mathbf{X}) = x_{1,\text{alarm}} - X_1 \leq 0)$, where the function $h(\mathbf{X})$ can be interpreted as a limit state function of its own. Thereby, the threshold value $x_{1,\text{alarm}}$ can be determined using any structural reliability method, as Equation (1) may be reformulated into the following equality, after taking the new information Z from the monitoring into account (Straub 2014, 2015):

$$p_{\text{F}|Z} = P(G(\mathbf{X}) \leq 0 \mid h(\mathbf{X}) \leq 0)$$
$$= \frac{P(\{G(\mathbf{X}) \leq 0\} \cap \{h(\mathbf{X}) \leq 0\})}{P(\{h(\mathbf{X}) \leq 0\})} = p_{\text{F,T}}, \quad (2)$$

where the numerator can be seen as parallel-system multiple failure modes and the denominator as a single failure mode. (In the simplest case with only one monitored parameter, Equation 2 is simply solved for the only unknown variable $x_{1,\text{alarm}}$). Here, $h(\mathbf{X})$ is defined so that critical behaviour corresponds to exceedance of the threshold, but thresholds against too low readings can also straightforwardly be determined. Note that if more than one threshold is to be established, Equation (1) becomes underdetermined, prompting additional assumptions or information (Spross and Gasch 2019). Such assumptions may, for example, concern the probability of exceeding the threshold. In this paper, however, we limit the study to one threshold.

2.2. Algorithm based on Subset simulation

The procedure to establish reliability-based thresholds from Equation (2) requires an iterative approach. This can be solved with crude Monte Carlo simulation (Spross and Johansson 2017), but when the limit state function requires input from a computationally demanding structural model (e.g. a Finite Element model), crude Monte Carlo becomes very inefficient. Spross and Gasch (2019), therefore, developed a general computational algorithm that efficiently computes such thresholds. The complete algorithm is provided in Appendix A and outlined in the following.

The algorithm applies Subset simulation (Au 2001, 2014), which is an adaptive form of Monte Carlo simulation that is particularly efficient for estimation of low failure probabilities when there are many random variables (Straub, Papaioannou, and Betz 2016). This efficiency facilitates the evaluation of the structural response with computationally demanding Finite Element models. A recent geotechnical application was presented by Gao et al. (2019).

In short, subset simulation calculates the p_{F} as a product of larger conditional probabilities by defining p_{F} from a number of nested intermediate failure events, such that $F_0 \supset F_1 \supset \ldots \supset F_M$, where F_0 is a certain event and F_M is the structural failure described by the event $\{G(\mathbf{X}) \leq 0\}$. The probability of failure can then be formulated as

$$p_{\text{F}} = P\left(\bigcap_{k=1}^{M} F_k\right) = \prod_{k=1}^{M} P(F_k | F_{k-1}), \quad (3)$$

in which $F_k = \{G(\mathbf{X}) \leq c_k\}$, where the limit c_k for each intermediate event corresponds to a predefined probability p_0 for the event F_k. The simulation entails a stepwise pushing of the sampling (by decreasing the value of c_k) toward the failure event of interest, where $c_M = 0$, corresponding to the investigated limit state. This is facilitated using Markov chain Monte Carlo (MCMC) simulation by letting the samples that satisfied the previous intermediate event F_{k-1} be the seeds of the next sampling round. Considering that p_0 is a predefined probability, the p_F can be estimated as

$$p_F \approx \hat{p}_F = p_0^{M-1} p_M, \qquad (4)$$

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

$$p_M = \frac{1}{N} \sum_{i=1}^{N} I_{F,M}(\mathbf{x_{M-1}}), \qquad (5)$$

where $I_{F,M}$ is the indicator function of F_M that is evaluated with the samples in the last round of sampling, $\mathbf{x}_{M-1}$, which were generated conditionally on the event F_{M-1}.

This general sampling procedure for subset simulation was adjusted in the algorithm for establishing reliability-based thresholds (Appendix A), to account for the information provided by the monitoring, by using the approach suggested by Straub, Papaioannou, and Betz (2016). To simulate the conditional probability of Equation (2), the intermediate events in Equation (3) are reformulated into $F_0^* \supset F_1^* \supset \ldots \supset F_M^*$, such that $F_k^* = \{G(\mathbf{X}) \leq c_k\} \cap \{h(\mathbf{X}) \leq 0\}$ for $k \geq 1$ and $F_0^* = \{h(\mathbf{X}) \leq 0\}$, which gives the calculated probability of attaining the limit state $G(\mathbf{X}) \leq 0$, conditional on the information Z that the threshold described by $h(\mathbf{X})$ has not been violated:

$$p_{F|Z} = P(G(\mathbf{X}) \leq 0 \mid h(\mathbf{X}) \leq 0) = P\left(\bigcap_{k=1}^{M} F_k^* \mid F_0^*\right)$$

$$= \prod_{k=1}^{M} P(F_k^* | F_{k-1}^*). \qquad (6)$$

Conceptually, the algorithm uses an iterative process to find the threshold $x_{1,alarm}$ that satisfies Equation (2), based on a first guess of $x_{1,alarm}$ (suggestions for reasonable starting values and revised guesses are provided in Spross and Gasch (2019)). In each iteration, the described Subset simulation procedure is performed using independent-component MCMC (Au and Wang 2014), aiming to minimise the error, $\hat{\varepsilon}$, between $p_{F|Z}$ and $p_{F,T}$. The iterations end when $\hat{\varepsilon}$ is less than a predefined tolerance, τ, which adjusts the accuracy of

how well the structural safety corresponds to $p_{F,T}$ when considering the information that the monitoring result is not violating $x_{1,alarm}$. In the practical application, the reliability-based threshold should be interpreted such that as long as $x_{1,alarm}$ has not been exceeded, the structural safety is acceptable with respect to the considered limit state.

2.3. Implementation details

In this study, all steps of the algorithm outlined in Section 2.2 and detailed in Appendix A are implemented in the general computational modelling software COMSOL Multiphysics (Comsol 2018). A crude sampling approach is adopted in all realisations of the random variables in $\mathbf{X}$. Important to emphasise is that our implementation puts no limitation on the structural model evaluated in steps 2b and 3dii of Table A.1 (in Appendix A), which can be 3-D or 2-D, and include, for example, material nonlinearity and time-dependent effects. Actually, the incorporation of our algorithm in a general Finite Element software gives us access to all functionality of the specific code when setting up a physical model. The algorithm is, in fact, not limited to structural cases, and could just as well be applied for monitoring the inflow of groundwater to a tunnel or the concentration of some chemical species in the groundwater close to the construction site. Any measurement error in the observation of X_1 can be considered straightforwardly by increasing its variability correspondingly in the algorithm when samples of X_1 are generated or evaluated from the numerical model. Details on models for measurement errors are provided by, for example, Baecher and Christian (2003). Generally, increasing measurement errors leads to stricter thresholds.

3. Practical application example

3.1. Description of the Stockholm bypass tunnel project

The Stockholm bypass project will provide a new route for the European highway E4 and will connect the Northern and Southern parts of Stockholm County with three lines in each direction, separated into two parallel 18-km-long tunnels. The tunnels are excavated in rock under lake Mälaren and the Lovön and Kungshatt islands west of Stockholm, Sweden (Figure 1). At the passage under lake Mälaren, south of Kungshatt, the excavation intersects a regional strike-slip fault zone, where the south side also has rotated in a reversed dip-slip movement. The fault zone is

Figure 1. Route of the southern part of the Stockholm Bypass. Published with permission from the Swedish Transport Administration.

approximately 200 m wide at the passage. The interpreted results from the pre-investigations indicate a number of parallel weakness zones mainly oriented along the fault line, but with occasional subhorizontal weakness zones in the tunnel direction, where blocks have rotated upwards. The pre-investigations indicate that the weakness zones, which consist of gauge, cataclasites, mylonites or breccias, may be limited in extension and surrounded by areas of higher quality rock mass.

When passing the regional fault zone, the tunnel will be located 64 m below the lake surface. The weak rock mass is overlain by till and clay, which forms the lake bottom. The engineering geological forecast identified two typical rock qualities, denoted as rock classes IV and V, within the fault zone. Rock class IV describes mainly the quality of the rock mass in the transition zone of the fault zone, while rock class V describes mainly the quality of the rock mass in the core of the fault zone. This paper analyses only rock class IV, which is characterised by an adjusted Rock Mass Rating (RMR) value within the range 38–59 with 40 being the assigned typical value. (For reference, rock class V has an adjusted RMR in the range 28–33 with 29 being the assigned typical value.) The Swedish "adjusted RMR" is used in early design phases; it is based on Bieniawski (1989), but sets ground water conditions to "completely dry" and discontinuity orientation to "very favourable", as these conditions are not yet known.

The tunnels are excavated with drilling and blasting and temporarily supported with a systematic pattern of rock bolts and a layer of shotcrete. Where poor rock quality is expected, a pipe umbrella system is installed as pre-support. Because of the poor rock mass quality and the large width of the tunnel, the excavation is divided into a sequence of gallery and bench excavations, with an advance rate of 2 m per sequence. As the expected rock cover is limited, extensive rock grouting is performed to limit the risk of stability issues caused by flowing ground conditions. A permanent cast concrete lining will be installed at a later stage, to account for long-term issues related to, for example, potential swelling clay. Further details of the geological conditions and the applied technical solution have been presented by Stille et al. (2019).

3.2. Analysed limit state and considered variables

3.2.1. Limit state definition

The excavation of the two tunnels is subject to several potential failure modes, including the collapse of the pipe umbrella system, failure of temporary support, face collapse and flowing ground. Instability caused by the attainment of the rock mass strength due to large in-situ stress is judged to be the main failure mechanism. This causes a loose core of rock that may lead to a progressive, large-scale collapse. A complicating factor

is the stress conditions, which are affected by the different advance rates of the two tunnel faces. The complete design is, therefore, rather complex.

For this reason, the case analysed in this paper is a simplification of the real tunnel design, to better highlight the features of an underground rock excavation that can be related to the reliability-based thresholds. We consider, therefore, only the failure mode related to the exceedance of the shotcrete compressive strength, as this is a good indicator of the structural behaviour of the main support system. Moreover, we limit the analysis to the cross section of one tunnel with a simplified geometry, without considering any effects from the other parallel tunnel. As illustrated in Figure 2, the tunnel is assumed to be subjected to an overburden of 27 m of rock with rock class IV, 26 m of soil and 11 m of lake water. The support consists of 300 mm of shotcrete and 5-m rock bolts. As we investigated potential roof collapse, we did not consider the pipe umbrella support system in the finite element model, as the main purpose of the pipe umbrellas is to prevent face collapse.

Shotcrete can fail due to either attained compression or attained tensile strength. In the general case, multiple failure modes are analysed. Here, however, we analyse only compressive failure, which gives the following limit state function:

$$G(\mathbf{X}) = f_c - \sigma, \tag{7}$$

where f_c is the compressive strength of the shotcrete and σ is the maximum compressive stress that *any* material point in the shotcrete support is subjected to; using the standard definition of stress in solid mechanics, we have that

$$\sigma = \max_{\mathbb{X} \in \Omega_s} [-\sigma_3(\mathbb{X})], \tag{8}$$

where σ_3 is the minimum (compressive) principal stress, $\mathbb{X}$ is the coordinate vector and Ω_s is the shotcrete domain. Thus, σ depends on the complex structural interaction between the rock support, the rock mass properties and the in-situ stresses in the rock.

This limit state corresponds to unsatisfactory serviceability due to the inward deformation of the shotcrete arch. Effects from bending deformation or local deformation around, for example, rock bolts, are not well described by this limit state, and rather imply a tensile failure of the shotcrete support. Extension of our procedure to consider also such failure modes is conceptually straightforward by considering multiple limit state functions.

3.2.2. Description of threshold violation as complementary limit state function
To verify structural safety, the vertical shotcrete displacement, δ, is to be monitored at the crown of the tunnel. Displacement monitoring is common in tunnel construction, especially when tunnelling in complex geological conditions with large geotechnical uncertainty, such as the fault zone passage under Lake Mälaren in the selected case. The observed displacements are used as an indirect indication of the attainment of the limit state in Equation (7), which therefore can be rewritten into

$$G(\mathbf{X}) = f_c - \delta C = 0, \tag{9}$$

where C is a transformation factor using observations of δ in the crown as indirect observations of σ anywhere in the shotcrete support. The C is given by

$$C = \frac{\sigma}{\delta} \Psi, \tag{10}$$

where Ψ is a model error describing the uncertainty in this transformation. The C can here be interpreted as an uncertain stiffness-like parameter of the shotcrete–rock interaction at the measurement point, which accounts for the behaviour of the entire support system. The magnitude and determination of C is further discussed in the next chapter.

By defining the event of identifying allowable shotcrete stress with the function

$$h(\mathbf{X}) = \delta_{\text{alarm}} - \delta \leq 0, \tag{11}$$

where δ_{alarm} is a threshold for unacceptable vertical crown displacement, Equation (2) can be applied to find the δ_{alarm} that must not be violated, in order to ensure that the $p_{\text{F,T}}$ is satisfied:

$$P(f_c - \delta C \leq 0 \mid \delta_{\text{alarm}} - \delta \leq 0) = p_{\text{F,T}}. \tag{12}$$

3.3. Application of subset simulation algorithm

3.3.1. Modelling of geotechnical uncertainty
To solve Equation (12) for δ_{alarm}, the subset simulation procedure described in Appendix A was applied. The structural interaction problem was analysed with the Finite Element Method. Using Spross and Gasch's (2019) notation, the basic random variables, $\mathbf{B} = [B_1, \ldots, B_n]$, considered in the Finite Element model were described by a joint probability density function (denoted $f_{\mathbf{B}}$ in Appendix A). Each marginal distribution of the basic random variables, B_i, is defined in accordance with Table 1. As rock mass deformation can be interpreted as a spatially averaging process, the marginal distributions in $\mathbf{B}$ reflect the
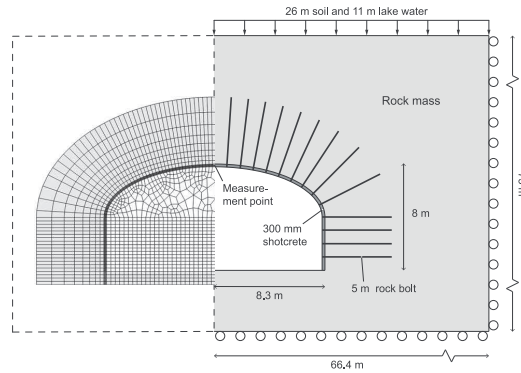
Figure 2. Analysed structural model of tunnel section. Left: Used mesh. Right: Geometries, boundaries, and loads.

epistemic uncertainty in the respective parameter's spatial average over the deformed volume. This is a reasonable assumption when the Finite Element model size is larger than a typical Representative Elementary Volume (REV) of a fractured rock mass (Esmaieli, Hadjigeorgiou, and Grenon 2010). Mean values of the spatial averages were assigned based on technical reports from the Stockholm bypass project, specifically for rock class IV. In lack of detailed statistical data on the ground conditions at the site, probability distribution types and coefficients of variations for the spatial averages were assigned based on the authors' judgement of reported pre-investigation results and design assumptions from the Stockholm bypass project, as well as available literature (e.g. JCSS 2001; Sari, Karpuz,

and Ayday 2010; Cai 2011; Langford and Diederichs 2013; Aladejare and Wang 2017). (In essence, the assigned variabilities are in line with what other researchers have used or reported, intending to illustrate a reasonable outcome of a more comprehensive Bayesian site characterisation. Thus, the coefficients of variation indicate the ranges within which the true spatial averages are expected to fall.) Material properties were assumed time-invariant. The distribution of Young's modulus of shotcrete, E_s, was assigned to reflect the aleatory uncertainty of a not-yet-sprayed concrete layer.

Geometries of the rock surface and the shotcrete thickness, as well as the shotcrete properties, are assumed to be without irregularities or spatial variability. While partly an effect of a need for mathematical

Table 1. Parameters considered in the analysed limit state. Mean values correspond to the characteristic values applied in the real-case for rock class IV and coefficients of variation (COV) are assigned to represent a realistic uncertainty of the spatial average over the modelled rock volume.

Parameter	Symbol	Marginal distribution type	Mean	COV
Rock mass				
Young's modulus of rock mass	E_r	Lognormal	3 GPa	0.15
Poisson's ratio of rock mass	ν_r	Lognormal	0.25	0.05
Density of overlying soil	ρ_s	Lognormal	2000 kg/m³	0.05
Density of rock mass	ρ_r	Lognormal	2650 kg/m³	0.05
Friction angle	φ	Lognormal	46°	0.05
Cohesion	c	Lognormal	1.0 MPa	0.25
In-situ stress perpendicular to tunnel[a]	σ_h	Lognormal	2.3 MPa	0.10
In-situ stress in tunnel direction[a]	σ_H	Lognormal	4.7 MPa	0.15
Shotcrete				
Compressive strength of shotcrete	f_c	Lognormal	25 MPa	0.15
Young's modulus of shotcrete	E_s	Lognormal	30 GPa	0.15
Poisson's ratio of shotcrete	ν_s		0.2	–
Young's modulus of rock bolt	E_b		210 GPa	–
Transformation factor[b]	C		3.51 GPa	0.135
			Triangular distribution parameters	
Confinement loss at support installation[c]	λ_{front}	Triangular	$a_{tr} = 0.49$; $b_{tr} = 0.80$; $c_{tr} = 0.53$	

[a]The σ_h and σ_H are based on in-situ stress measurements in the Stockholm region. The randomly generated values of σ_h and σ_H are increased with respect to the depth below the rock surface by a constant 13.75 and 37.5 kPa/m, respectively.
[b]Determined as described in the section estimation of transformation model error.
[c]Determined as described in the section estimation of confinement loss at support installation.

convenience in this study, the assumption is judged to be a reasonable simplification for the analysed large-scale serviceability limit state of a rather thick lining (300 mm) at the design stage. We base this on recent research studies on the effect of spatial variability of shotcrete properties and thickness on its failure behaviour (Bjureland et al. 2020; Sjölander, Ansell, and Malm 2021). They indicate that even for local block-fall failures, there is a substantial averaging effect. Furthermore, Malmgren and Nordlund (2008) studied numerically the effect of having an uneven rock surface and an evenly thick shotcrete lining. For this case, they found that the surface roughness is an important factor for the shotcrete behaviour, as it inflicted bending of the shotcrete at the edges of the rock surface in their model. In reality, however, rock surface irregularities tend to be evened out in the spraying of the shotcrete, especially for thicker linings, reducing the bending. (For ultimate limit states or thin linings, the effect of irregularities in the rock surface and shotcrete thickness may be more prominent and therefore require that spatial variation be accounted for in the model, to capture the effect of local weaknesses.)

3.3.2. Set-up of structural model

Dimensions of the simplified version of the tunnel used to set up our finite element model are shown in Figure 2. The rock bolts are placed with a spacing of 1 m, and
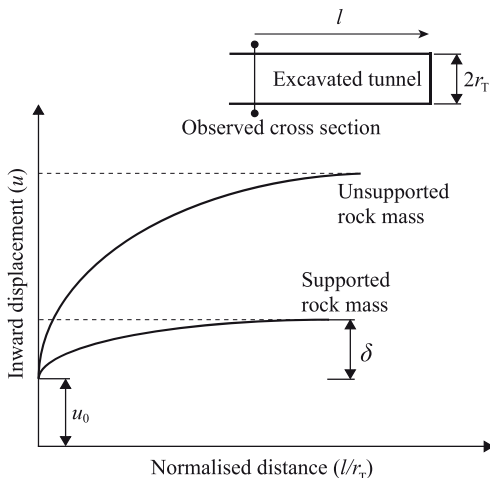

Figure 3. Principal behaviour of inward tunnel displacement as a tunnel is excavated. The u denotes the total displacement of the rock mass, while δ denotes the displacement of the shotcrete (Bjureland et al. (2017), CC-BY-NC-ND 4.0, https://creativecommons.org/licenses/by-nc-nd/4.0/, with minor adjustments).

relevant boundary conditions and loads are depicted in the figure. Only one-half of the tunnel-rock geometry was included in the model by assuming a vertical symmetry plane. The geometry was discretised with a mesh as shown in the left part of the figure (only showing the mesh close to the tunnel) and with an assumption of 2-D plane strain. In total, the mesh consists of 3123 elements, and, given a quadratic displacement field, the finite element model includes 19350 degrees of freedom (DOFs). Note that the rock bolts are placed along mesh element boundaries and, thus, are assumed to fully interact with the rock/shotcrete, and therefore add no additional DOFs to the model.

Both rock bolts and the shotcrete were assumed to be elastic, while the rock mass was considered elastoplastic. A Dücker–Prager yield criterion with the non-associated flow was used for the rock mass, with parameters determined from the Mohr–Coloumb criterion. The elastic constants of the rock mass and shotcrete, the cohesion, c, and the friction angle, φ, of the rock mass were considered random variables, while the dilatancy angle, ψ, controlling the plastic flow in the rock mass was assumed fully correlated with the friction angle, such that $\psi = \varphi/8$.

In 2-D modelling of rock tunnels, a key parameter is the inward displacement of the rock mass, u, here taken as the vertical displacement at the measurement point indicated in Figure 2. The principal behaviour of this displacement and its relation to δ are shown in Figure 3: a supported rock mass in a considered tunnel section exhibits smaller displacement than an unsupported rock mass, as the excavation continues. A design challenge lies in the assessment of how much inward displacement, u_0, and corresponding confinement loss, λ_{front}, has already occurred in front of the tunnel face. This structural behaviour can be analysed with longitudinal displacement profiles (LDP). However, generating LDPs for the general case requires full 3-D models to capture effects from, for example, non-hydrostatic stress conditions and complex tunnel geometries (Vlachopoulos and Diederichs 2009; Langford and Diederichs 2013). Because of the substantial computational effort that this would require in a reliability analysis, we have applied a simplified approach in 2-D to evaluate probabilistically the u_0 and the corresponding confinement loss, using findings of Vlachopoulos and Diederichs (2009) as described in the following.

In our 2-D model, the confinement loss is considered by using an incremental analysis, where the support given by the rock to be excavated is incrementally removed by reducing its stiffness and stress. This means that the tunnel was actually included in the discretised geometry, as is shown in the left part of Figure
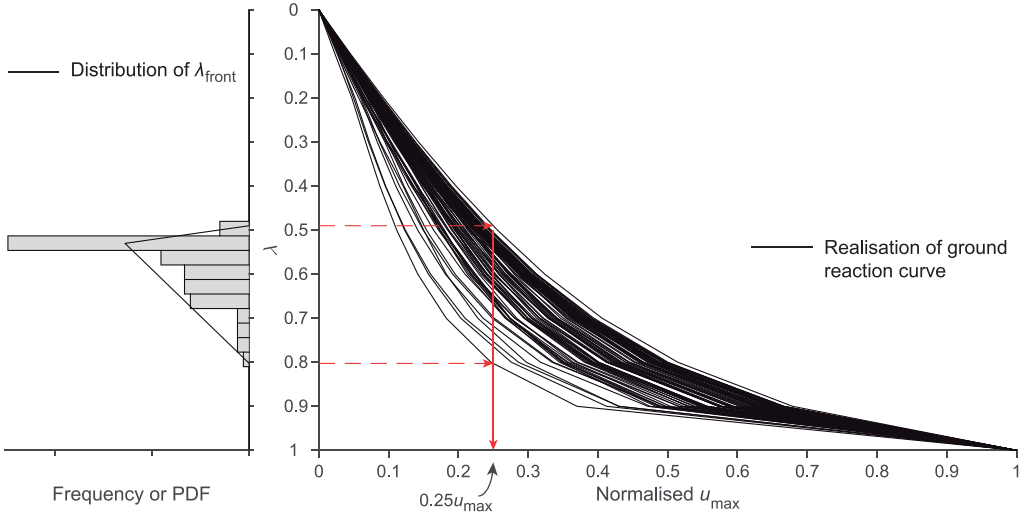
Figure 4. Estimation of triangular probability density function for the confinement loss λ_{front} at 25% of u_{max}, based on 100 realisations of the excavation in the 2-D model.

2. This incremental procedure can be described by the confinement loss parameter, λ, which ranges from $\lambda = 0$ simulating full confinement to $\lambda = 1$ simulating no confinement (González-Nicieza et al. 2008). Formally, the stress tensor $\boldsymbol{\sigma}$ and the fourth-order elasticity tensor $\mathbf{D}$ in the domain to be excavated can be described by

$$\sigma = (1 - \lambda)\sigma_0, \tag{13a}$$

$$\mathbf{D} = (1 - \lambda)\mathbf{D}_0, \tag{13b}$$

where σ_0 is the in-situ stress, and $\mathbf{D}_0$ describes the stiffness of the intact rock. By incremental increase of λ toward no confinement in the model, the resulting u can be plotted as a ground reaction curve.

Moreover, the support structure (shotcrete and bolts) is installed in a stress-free state at an intermediate step, where the support from the excavated rock has been partially removed. Assuming that the rock support is installed at the tunnel front, the λ_{front} describes the point in the incremental relaxation of the applied pressure in the 2-D model, when the support is to be implemented in the model. The λ_{front} was considered a random variable and its determination is further discussed in the next section. To summarise, the FE simulation can be described by the following steps:

1. Initialise in-situ stress in the non-excavated rock domain.

2. Incrementally remove the support from the tunnel by reducing its stiffness and stress (Equation (13)).
3. Insert the rock support system in a stress-free state given by λ_{front}.
4. Continue incremental removal procedure until $\lambda = 1$.

3.3.3. Estimation of confinement loss at support installation

To estimate the displacement u_0 at the tunnel front, we assessed the ratio of the maximum radius of the plastic zone and the approximate equivalent tunnel radius (r_P/r_T) to be approximately 2, based on initial simulations of a 2-D model of the analysed tunnel section (unsupported). According to Vlachopoulos and Diederichs (2009), $r_P/r_T = 2$ corresponds to the u_0 being 25% of the expected maximum deformation of the unsupported tunnel, u_{max}. For 100 simulated ground reaction curves of the 2-D model without support, using crude Monte Carlo, we then investigated the uncertainty in the confinement loss that corresponded to $u_0 = 0.25u_{\text{max}}$, that is, the confinement loss λ_{front} (Figure 4). Based on the 100 observations of λ_{front}, we assigned it a triangular distribution, as presented in the left chart of Figure 4 and Table 1.

3.3.4. Estimation of transformation model error

The magnitude of the transformation model error Ψ in Equation (10) was evaluated from the crude Monte

Table 2. Simulation constants used in the example.

Constant	Value
$p_{F,T}$	0.005
p_0	0.1
T	0.1
$N \cdot \kappa$	2500·3

Carlo simulations of the structural model (step 2.b in Table A.1 in Appendix A), by calculating the variability in the ratio σ/δ, based on the output from the finite element analyses. The variability of Ψ is shown in Table 1 as the COV of C. The Ψ was then implemented in the limit state evaluations in the subsequent steps 2.c and 3.d.ii in the algorithm.

3.3.5. Simulation constants

To run the algorithm that determines δ_{alarm} for the analysed case, the following simulation constants need to be defined: target probability of failure, $p_{F,T}$; the probability of the intermediate events in the subset simulation, p_0; the tolerance, τ, of the error between $p_{F|Z}$ and $p_{F,T}$ for the proposed threshold; the number of samples in the subset simulation, N; and the constant κ, which ensures that there are enough samples in the initial crude Monte Carlo simulation (see details in Appendix A).

As summarised in Table 2, the $p_{F,T}$ was set to 0.005, as the analysed limit state can be seen as a serviceability limit state for the temporary rock support. The $p_{F,T}$ used corresponds to suggested target probabilities for other serviceability limit states (Fenton, Naghibi, and Griffiths 2016), but in practice the required safety level is an issue to study further in development of

reliability-based design codes. The p_0 was set to 0.1, following recommendations by Zuev et al. (2012), and τ was set to 0.1. Based on preliminary runs on the algorithm, κ was set to 3.

3.4. Calculation results

Figure 5 shows the simulation results for the last, accepted threshold, $\delta_{alarm} = 4.4$ mm. Figure 5(a) shows an output space in terms of deformation quantities, while Figure 5(b) shows the same simulations but in terms of strength–stress. The grey dots and the contour lines represent the initial crude Monte Carlo simulations (step 2 in the algorithm in Appendix A). The black dots represent the accepted data points after truncation at the proposed threshold value (step 3b). The red dots represent the last subset simulation level, that is, the simulations used to determine the probability of event F_M^* (cf. Equation (6) and step 3d). The probability of satisfying the threshold and therefore also satisfying the $p_{F,T}$ is 71%. Without adhering to an established threshold, the calculated p_F would be 0.012 for the analysed design solution in this limit state in rock class IV. This illustrates the advantage of designing with planned monitoring in an observational approach: if the monitoring were not accounted for in the structural safety assessment, a more conservative design would be required to satisfy $p_{F,T} = 0.005$.

Figure 6 illustrates four evaluations of the final element model (marked in Figure 5), where (a) represents a typical structural behaviour, (b) represents
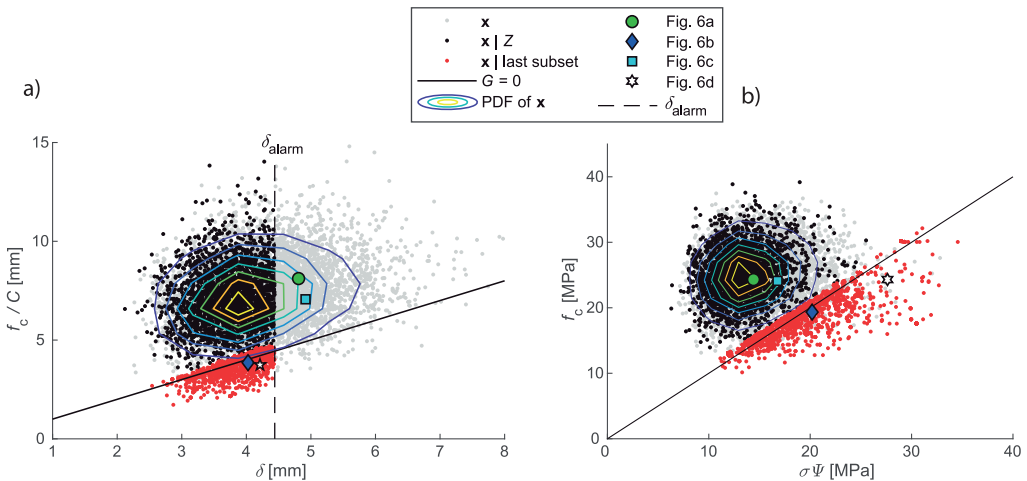

Figure 5. Simulation results: (a) Output space in deformation quantities. (b) Output space in stress–strength quantities after transformation with C. The finite element model evaluations of four specific simulations are shown in Figure 6.
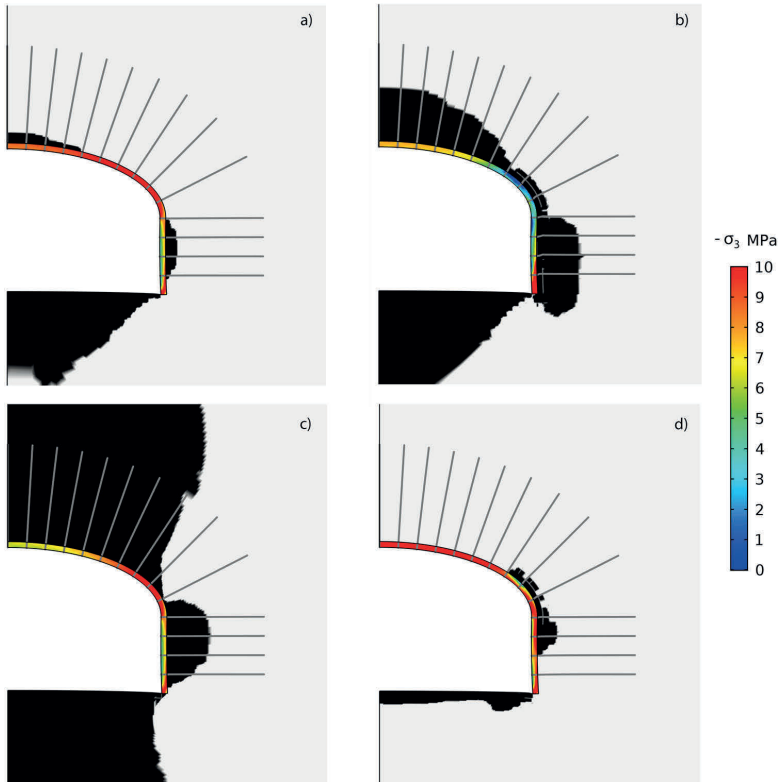
Figure 6. Simulation results capturing different structural behaviour in the FE-model. Black areas indicate the plasticised rock mass around the tunnel. The simulations are marked in Figure 5.

failure behaviour (i.e. violation of the limit state function), (c) represents a simulation with low horizontal in-situ stress that however did not violate the limit state and (d) represents a failure behaviour despite limited plasticised rock mass volume. The difference in structural behaviour between the figures indicates the complexity of the case: there are many possible combinations of input data, and the resulting behaviour in the evaluation of the structural model in each simulation is very difficult to predict in advance.

4. Discussion

4.1. Effects of model errors and other simplifications

Model errors affect the calculated threshold considerably for this application. In the presented example, we introduced a transformation model error (Ψ) to account for the uncertainty related to the fact that we used the displacement measurement in the tunnel crown to capture f_c exceedance anywhere in the lining. A comparison of the two graphs in Figure 5 illustrates the significance of this error clearly: the data points representing displacements below the threshold in Figure 5(a) (black dots) exhibit a considerable scatter in the strength–stress output plot of Figure 5(b). The uncertainty represented by Ψ depends partly on the extent of the monitoring programme. By measuring the shotcrete displacement at the expected location of the largest σ, instead of at the tunnel crown, a smaller transformation error would be expected.

Another model error is introduced through our simplified approach to assess the occurred confinement loss at the tunnel front at the time of support installation (λ_{front} in Figure 4). In this calculation example, we derived a triangular distribution for this model error. For simplicity, we did not, however, account for any uncertainty in the estimation of the ratios r_P/r_T and u_0/u_{max} that were needed to assess the uncertainty of λ_{front}. A straightforward approach to this issue would be to also assign a model uncertainty to u_0/u_{max}, for
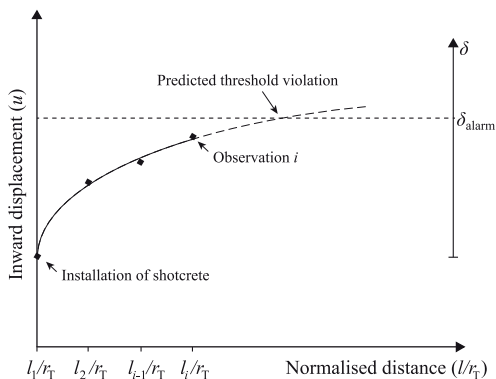
Figure 7. Expected logarithmic inward displacement behaviour with increased distance to the tunnel front. Using Bayesian regression analysis on the retrieved data points, a prediction can be made of the potential violation of the threshold.

example, by assigning this ratio a uniform distribution between, say, 0.2 and 0.3, before evaluating λ_{front}, instead of using the deterministic value 0.25.

We also disregarded the potential correlations between λ_{front}, the rock mass properties, and the in-situ stress conditions; ground conditions that are prone to cause substantial wall deformation (large u_{max}), because of significant plastic ground behaviour, may potentially be associated with more curved shapes of the normalised ground reaction curve (Figure 4). Such curved shapes may, in turn, be associated with less remaining confinement at the tunnel front (i.e. larger λ_{front}).

The calculation example assumes that the shotcrete is installed directly at the tunnel front and achieves full strength instantly. In practice, the tunnel support would be installed normally some distance behind the tunnel front and gain strength as the shotcrete cures. The principles of how to account for these aspects in tunnel support design have been discussed for deterministic design by, for example, Chang (1994), Carranza-Torres and Fairhurst (2000) and Oke, Vlachopoulos, and Diederichs (2018), but their application to reliability-based design remains an issue for future studies.

In addition to the aforementioned model errors, there may also be intrinsic errors in the applied FE-model. Such model errors have not been quantified and considered in this paper, which imposes a degree of arbitrariness on the calculated threshold. We note however that the effect of intrinsic model errors remains a general issue for future research in practical application of reliability-based analyses of rock engineering structures.

4.2. Accuracy of threshold and challenges with large uncertainties

The iterative nature of the applied algorithm (Appendix A) implies that a proposed threshold is accepted, if the conditional event of not violating the proposed threshold provides a calculated p_F that is close enough to $p_{F,T}$. This tolerance (τ) between p_F and $p_{F,T}$ is defined by the user. The larger the τ is, the faster the algorithm is expected to find an acceptable threshold value, but the larger becomes the potential deviation from the required $p_{F,T}$ when the threshold is applied. We believe a quite large τ can often be acceptable in practice, especially for serviceability limit states, meaning that the established threshold only needs to provide alarms related to the correct order of magnitude of the required $p_{F,T}$.

In case there is only limited information regarding the site conditions, the uncertainty of the involved geotechnical parameters may be considerable. While this does not pose a problem for the algorithm itself, its accuracy might be impaired for cases where a non-linear FE-model is considered. With large uncertainty represented in the probability distributions, the likelihood increases of drawing sample combinations that characterise extremely low-quality rock mass with virtually no strength. For such samples, the chosen structural model may not converge for a non-negligible number of evaluations, that is, decreasing the number of valid data points. This would indicate a need for another approach for discretisation of the real conditions, for example, by using another material model or numerical method. Such considerations would, of course, also be needed in a deterministic sensitivity analysis for the same situation. For practical application, we recommend to monitor the number of non-converged model evaluations to ensure sufficient accuracy of the evaluated threshold.

4.3. Practical implementation with the observational method

The definition of the threshold (Equation (1)) implies that a violated threshold by definition means exceeding the $p_{F,T}$. For the analysed limit state, the shotcrete in a considered cross section is loaded sequentially as the tunnel excavation progresses with each blasting round. Measuring the occurred inward displacement of the shotcrete (δ) after each blasting round, a logarithmic pattern can be expected between δ and the normalised distance to the tunnel front, l/r_T (Figure 7). After a few blasting rounds with corresponding displacement measurements at distances $[l_1, l_2, \ldots, l_i]$ from the tunnel front, a prediction of the final displacement can be

made, using a linear regression model with logarithmic transformation on the format

$$\delta = \alpha + \beta \ln\left(\frac{l}{r_{\mathrm{T}}}\right), \tag{14}$$

where α and β are regression coefficients. Based on initial work by Holmberg and Stille (2009), Bjureland et al. (2017) showed how such measurement data can be analysed and implemented within a reliability-based framework for the observational method: using Equation (14) in a Bayesian framework, the final displacement can be predicted through extrapolation, which will be increasingly accurate as more data is collected. If this prediction indicates that the derived threshold δ_{alarm} is likely to be violated, the responsible decision maker regarding the tunnel support is alerted to take action, for example, by ensuring that additional support is installed. The use of reliability-based thresholds as a part of the observational method can thereby integrate the shotcrete design into a larger risk-based framework for tunnel design and construction, which was proposed by Bjureland et al. (2020).

In our application example, we considered only one measurement point in the cross section and let the transformation factor C account for the behaviour of the complete shotcreted area. However, the roof and the walls are typically subjected to different structural behaviour, and moreover, the structural behaviour is affected considerably by the elastoplastic behaviour of the rock mass, as indicated in Figure 6. It is, therefore, likely to be necessary in a real-world application to implement more than one point of measurement and to consider multiple limit state functions, such as both tensile and compressive failure. This requires several thresholds to be derived. However, as previously mentioned, having more than one threshold makes Equation (1) underdetermined. How to find the *optimal* set of threshold levels in $\mathbf{x}_{\mathrm{alarm}}$ remains an issue for future research.

5. Conclusion

We have presented an extensive calculation example of how reliability-based thresholds can be established for the inward displacement of a shotcrete lining in a rock tunnel. The example considered a serviceability limit state function, for which a finite element model was needed to evaluate the ground–structure interaction associated with the loss of confinement from the advancing tunnel and the installed support. Because of the considerable computational effort required to evaluate the final element model probabilistically, an algorithm using subset simulation was employed. The effect of model uncertainty related to the confinement loss and transformation uncertainty related to information from indirect monitoring was investigated and discussed extensively. In particular, we find that there is a need for future research on how to consider the uncertainty in the modelling of confinement loss in 2-D models, as full longitudinal displacement profiles (LDPs) in 3-D still may be unreasonable to employ in probabilistic analyses.

The derived threshold may serve as a key component in designing a tunnel with the observational method. Our analysis in this paper showed how applying such thresholds could facilitate less conservative designs, as the thresholds ensure sufficient safety by being related to a target reliability and account for prevailing uncertainties. Thereby, the threshold can become an integrated part in the performed risk management work in the tunnel project.

ORCID

Johan Spross http://orcid.org/0000-0001-5372-7519
Tobias Gasch http://orcid.org/0000-0002-8000-6781
Fredrik Johansson http://orcid.org/0000-0002-8152-6092

References

Aladejare, A. E., and Y. Wang. 2017. "Evaluation of Rock Property Variability." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 11 (1): 22–41.

Au, S.-K., and J. L. Beck. 2001. "Estimation of Small Failure Probabilities in High Dimensions by Subset Simulation." *Probabilistic Engineering Mechanics* 16 (4): 263–277.

Au, S.-K., and Y. Wang. 2014. *Engineering Risk Assessment with Subset Simulation*. Singapore: John Wiley & Sons.

Baecher, G. B., and J. T. Christian. 2003. *Reliability and Statistics in Geotechnical Engineering*. Chichester: John Wiley & Sons.

Bieniawski, Z. T. 1989. *Engineering Rock Mass Classifications: A Complete Manual for Engineers and Geologists in Mining, Civil, and Petroleum Engineering*. New York, NY: John Wiley & Sons.

Bjureland, W., F. Johansson, A. Sjölander, J. Spross, and S. Larsson. 2019. "Probability Distributions of Shotcrete Parameters for Reliability-based Analyses of Rock Tunnel Support." *Tunnelling and Underground Space Technology* 87: 15–26.

Bjureland, W., F. Johansson, J. Spross, and S. Larsson. 2020. "Influence of Spatially Varying Thickness on Load-bearing Capacity of Shotcrete." *Tunnelling and Underground Space Technology* 98 (4): 103336.

Bjureland, W., J. Spross, F. Johansson, A. Prästings, and S. Larsson. 2017. "Reliability Aspects of Rock Tunnel Design with the Observational Method." *International Journal of Rock Mechanics and Mining Sciences* 98: 102–110.

Cai, M. 2011. "Rock Mass Characterization and Rock Property Variability Considerations for Tunnel and Cavern Design." *Rock Mechanics and Rock Engineering* 44 (4): 379–399.

Carranza-Torres, C., and C. Fairhurst. 2000. "Application of the Convergence-confinement Method of Tunnel Design to Rock Masses That Satisfy the Hoek-Brown Failure Criterion." *Tunnelling and Underground Space Technology* 15 (2): 187–213.

CEN. 2004. *En 1997-1:2004 Eurocode 7: Geotechnical Design – Part 1: General Rules*. Brussels: European Committee for Standardisation.

Chang, Y. 1994. *Tunnel Support with Shotcrete in Weak Rock: A Rock Mechanics Study*. Stockholm: KTH Royal Institute of Technology.

Comsol. 2018. *Comsol Multiphysics® Version 5.4*. Stockholm: Comsol AB.

Contreras, L.-F., and E. T. Brown. 2019. "Slope Reliability and Back Analysis of Failure with Geotechnical Parameters Estimated Using Bayesian Inference." *Journal of Rock Mechanics and Geotechnical Engineering* 11 (3): 628–643.

Esmaieli, K., J. Hadjigeorgiou, and M. Grenon. 2010. "Estimating Geometrical and Mechanical Rev Based on Synthetic Rock Mass Models at Brunswick Mine." *International Journal of Rock Mechanics and Mining Sciences* 47 (6): 915–926.

Feng, X., and R. Jimenez. 2015. "Estimation of Deformation Modulus of Rock Masses Based on Bayesian Model Selection and Bayesian Updating Approach." *Engineering Geology* 199 (12): 19–27.

Feng, X., R. Jimenez, P. Zeng, and S. Senent. 2019. "Prediction of Time-dependent Tunnel Convergences Using a Bayesian Updating Approach." *Tunnelling and Underground Space Technology* 94: 103118.

Fenton, G. A., F. Naghibi, and D. V. Griffiths. 2016. "On a Unified Theory for Reliability-based Geotechnical Design." *Computers and Geotechnics* 78: 110–122.

Frank, R., C. Bauduin, R. Driscoll, M. Kavvadas, N. Krebs Ovesen, T. Orr, and B. Schuppener. 2004. *Designers' Guide to En 1997-1 Eurocode 7: Geotechnical Design-general Rules*. London: Thomas Telford.

Gao, G.-H., D.-Q. Li, Z.-J. Cao, Y. Wang, and L. Zhang. 2019. "Full Probabilistic Design of Earth Retaining Structures Using Generalized Subset Simulation." *Computers and Geotechnics* 112 (8): 159–172.

González-Nicieza, C., A. E. Álvarez-Vigil, A. Menéndez-Díaz, and C. González-Palacio. 2008. "Influence of the Depth and Shape of a Tunnel in the Application of the Convergence–Confinement Method." *Tunnelling and Underground Space Technology* 23 (1): 25–37.

Holmberg, M., and H. Stille. 2009. "Observationsmetoden Och Deformationsmätningar Vid Tunnelbyggande [the Observational Method and Deformation Measurements in Tunnels]." Report 93. Stockholm: BeFo.

JCSS. 2001. "Probabilistic Model Code." Joint Committee on Structural Safety.

Kroetz, H. M., N. A. Do, D. Dias, and A. T. Beck. 2018. "Reliability of Tunnel Lining Design Using the Hyperstatic Reaction Method." *Tunnelling and Underground Space Technology* 77: 59–67.

Langford, J. C., and M. S. Diederichs. 2013. "Reliability Based Approach to Tunnel Lining Design Using a Modified Point Estimate Method." *International Journal of Rock Mechanics and Mining Sciences* 60: 263–276.

Li, X. Y., L. M. Zhang, S. H. Jiang, D. Q. Li, and C. B. Zhou. 2016. "Assessment of Slope Stability in the Monitoring Parameter Space." *Journal of Geotechnical and Geoenvironmental Engineering* 142 (7): 04016029.

Lü, Q., Z.-P. Xiao, J. Ji, and J. Zheng. 2017. "Reliability Based Design Optimization for a Rock Tunnel Support System with Multiple Failure Modes Using Response Surface Method." *Tunnelling and Underground Space Technology* 70: 1–10.

Malmgren, L., and E. Nordlund. 2008. "Interaction of Shotcrete with Rock and Rock Bolts—A Numerical Study." *International Journal of Rock Mechanics and Mining Sciences* 45 (4): 538–553.

Napa-García, G. F., A. T. Beck, and T. B. Celestino. 2018. "Risk Analysis of Fractured Rock Mass Underground Structures." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 12 (2): 123–134.

Oke, J., N. Vlachopoulos, and M. Diederichs. 2018. "Improvement to the Convergence–Confinement Method: Inclusion of Support Installation Proximity and Stiffness." *Rock Mechanics and Rock Engineering* 51 (5): 1495–1519.

Palmström, A., and H. Stille. 2015. *Rock Engineering*. London: Thomas Telford.

Peck, R. B. 1969. "Advantages and Limitations of the Observational Method in Applied Soil Mechanics." *Géotechnique* 19 (2): 171–187.

Sari, M., C. Karpuz, and C. Ayday. 2010. "Estimating Rock Mass Properties Using Monte Carlo Simulation: Ankara Andesites." *Computers & Geosciences* 36 (7): 959–969.

Schubert, W. 2008. "The Development of the Observational Method." *Geomechanics and Tunnelling* 1 (5): 352–357.

Schweckendiek, T., and A. C. W. M. Vrouwenvelder. 2013. "Reliability Updating and Decision Analysis for Head Monitoring of Levees." *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 7 (2): 110–121.

Sjölander, A., A. Ansell, and R. Malm. 2021. "Variations in Rock Support Capacity Due to Local Variations in Bond Strength and Shotcrete Thickness." *Engineering Failure Analysis* 128: 105612.

Spross, J., and T. Gasch. 2019. "Reliability-based Alarm Thresholds for Structures Analysed with the Finite Element Method." *Structural Safety* 76: 174–183.

Spross, J., and F. Johansson. 2017. "When Is the Observational Method in Geotechnical Engineering Favourable?" *Structural Safety* 66: 17–26.

Spross, J., and S. Larsson. 2014. "On the Observational Method for Groundwater Control in the Northern Link Tunnel Project, Stockholm, Sweden." *Bulletin of Engineering Geology and the Environment* 73 (2): 401–408.

Spross, J., H. Stille, F. Johansson, and A. Palmstrøm. 2018. "On the Need for a Risk-based Framework in Eurocode 7 to Facilitate Design of Underground Openings in Rock." *Rock Mechanics and Rock Engineering* 51 (8): 2427–2431.

Spross, J., H. Stille, F. Johansson, and A. Palmstrøm. 2020. "Principles of Risk-based Rock Engineering Design." *Rock Mechanics and Rock Engineering* 53: 1129–1143.

Stille, H., and M. Holmberg. 2010. "Examples of Applications of Observational Method in Tunnelling." *Geomechanics and Tunnelling* 3 (1): 77–82.

Stille, B., F. Johansson, F. R. Bayona, R. B. Estrada, and M. Roslin. 2019. "Stockholm Bypass Project – Passage Under the Lake Mälaren." In *Proceedings of the WTC 2019 ITA-AITES World Tunnel Congress*, edited by D. Peila, G. Viggiani, and T. Celestino, 1569–1578. London: CRC Press.

Straub, D. 2014. "Value of Information Analysis with Structural Reliability Methods." *Structural Safety* 49: 75–85.

Straub, D., and I. Papaioannou. 2015. "Bayesian Updating with Structural Reliability Methods." *Journal of Engineering Mechanics* 141 (3): 04014134.

Straub, D., I. Papaioannou, and W. Betz. 2016. "Bayesian Analysis of Rare Events." *Journal of Computational Physics* 314: 538–556.

Vlachopoulos, N., and M. S. Diederichs. 2009. "Improved Longitudinal Displacement Profiles for Convergence Confinement Analysis of Deep Tunnels." *Rock Mechanics and Rock Engineering* 42 (2): 131–146.

Zuev, K. M., J. L. Beck, S.-K. Au, and L. S. Katafygiotis. 2012. "Bayesian Post-processor and Other Enhancements of Subset Simulation for Estimating Failure Probabilities in High Dimensions." *Computers & Structures* 92-93: 283–296.

Appendix A

The applied algorithm for establishment of reliability-based thresholds is presented in Table A.1. The purpose of each step 1–4 is described in the following:

Step 1 contains definition of the simulation constants (Table 2), the limit state function $G(\mathbf{X})$ and the basic random variables in the joint probability density function $f_\mathbf{B}$. (The basic random variables in the vector $\mathbf{B}$ are used as input to the structural model, while $\mathbf{X}$ strictly contains the random parameters of the limit state function.)

Step 2 performs an initial round of crude Monte Carlo simulation, corresponding to the event F_0 in Eq. (3), by generating random samples $\mathbf{b}$ from $\mathbf{B}$ (i.e. random variables in Table 1) and applying them to the FE-model, which provides the output ($\mathbf{x}$) necessary to evaluate $G(\mathbf{x})$. A first guess of the

threshold $x_{1,\mathrm{alarm}}^{(1)}$ is made, corresponding to that of δ_{alarm} in the case study. Here, we follow Spross and Gasch's (2019) suggestion and assign based on the initial simulation $\delta_{\mathrm{alarm}}^{(1)} = (\max \boldsymbol{\delta} - \bar{\boldsymbol{\delta}})/2$, where $\boldsymbol{\delta}$ is a vector of all simulated displacements at the tunnel crown in the finite element model and $\bar{\delta}$ is their mean.

Step 3 is an iterative loop searching for an acceptable threshold, which ends when the error $\hat{\varepsilon}$ between the calculated $\hat{p}_F$ and $p_{F,T}$ is less than the tolerance τ. To satisfy the conditional event F_0^*, only simulated responses that are not violating the present threshold suggestion, $x_{1,\mathrm{alarm}}^{(j)}$, are retained. This ensures that the $p_{F,T}$ of the structure is satisfied as long as the threshold is not violated. After this, the Subset simulation with independent-component MCMC (Steps 3.c–e) is initiated to evaluate efficiently the probability of failure, $\hat{p}_F$, given non-violation of the present threshold suggestion, as described by Equation (6). When the error $\hat{\varepsilon}$ corresponding to the present threshold suggestion has been determined, a new suggestion $x_{1,\mathrm{alarm}}^{(j+1)}$ is made (Step 3.g), striving to reduce the range of possible remaining threshold suggestions given the discarded previous suggestions. Here, we follow Spross and Gasch's (2019) simple suggestion and set $x_{1,\mathrm{alarm}}^{(j+1)}$ to be the average value of the maximum and minimum values of the remaining possible range.

When $\hat{\varepsilon} < \tau$, Step 4 displays the accepted reliability-based threshold.

Table A.1. Applied algorithm for establishment of reliability-based thresholds. (Spross and Gasch (2019) with minor adjustments, CC-BY 4.0, https://creativecommons.org/licenses/by/4.0/).

1. Definition of simulation constants and basic, case-specific data.

    a. Select $p_{F,T}$, N, p_0, τ and κ (κ is a factor to provide sufficient numbers 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 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 threshold, $x_{1,\mathrm{alarm}}^{(1)}$.
    e. $j = 1$.

3. While $\hat{\varepsilon} > \tau$ (iterative loop to find $x_{1,\mathrm{alarm}}$)
    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 $x_1 < x_{1,\mathrm{alarm}}^{(j)}$, into matrix $\mathbf{x}_Z$. If the number of acceptable sample sets is less than N, increase κ 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 $\mathbf{x}_k^{(\mathrm{seed})}$, while $\mathbf{b}_k^{(\mathrm{seed})}$ 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 $N_s = 1/p_{0(N_s)}$ sets of conditional samples $\tilde{\mathbf{b}}_i = [\tilde{b}_i^{(1)}, \tilde{b}_i^{(l)}, \ldots, \tilde{b}_i^{(l)}]$ from a proposal PDF $q_i^*(b_i^{(\mathrm{seed})})$ for each basic parameter in $\mathbf{B}$.
            – Calculate for each proposal $r_i^{(l)} = \frac{q_i^*(b_i^{(\mathrm{seed})}; \tilde{b}_i^{(l)}) f_{B_i}(\tilde{b}_i^{(l)})}{q_i^*(\tilde{b}_i^{(l)}; b_i^{(\mathrm{seed})}) f_{B_i}(b_i^{(\mathrm{seed})})}$.
            – Generate $u_i^{(l)}$ uniformly distributed on [0,1] for each proposal.

– Set for all l and all basic parameters,

$$\bar{b}_i^{(l)} = \begin{cases} \tilde{b}_i^{(l)} & \text{if } u_i^{(l)} < r_i^{(l)} \\ b_i^{(\text{seed})} & \text{otherwise} \end{cases} \quad \text{and collect them in } \bar{\mathbf{b}}_k.$$

ii. Run the structural model for each set of samples in $\bar{\mathbf{b}}_k$ to find the complete set of proposed samples, $\bar{\mathbf{x}}_k$

iii. Set, for all N parameter sets of proposed samples,

$$\mathbf{x}_k = \begin{cases} \bar{\mathbf{x}}_k & \text{if } \bar{\mathbf{x}}_k \in F_k^* \\ \mathbf{x}_k^{(\text{seed})} & \text{otherwise} \end{cases}.$$

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 $\hat{p}_F = p_0^{k-1} \dfrac{N_F}{N}$.

f. Calculate the error between $\hat{p}_F$ and $p_{F,T}$ as $\hat{\varepsilon} = \dfrac{|(\hat{p}_F - p_{F,T})|}{p_{F,T}}$.

g. Set $x_{1,\text{alarm}}^{(j+1)}$ to be $\begin{cases} < x_{1,\text{alarm}}^{(j)} & \text{if } \hat{p}_F > p_{F,T}(1 + \tau) \\ > x_{1,\text{alarm}}^{(j)} & \text{if } \hat{p}_F < p_{F,T}(1 - \tau) \end{cases}$.

h. $j = j + 1$.

4. $x_{1,\text{alarm}} = x_{1,\text{alarm}}^{(j-1)}$.

BeFo