Proceedings of the XVII ECSMGE-2019
Geotechnical Engineering foundation of the future
ISBN 978-9935-9436-1-3

doi: 10.32075/17ECSMGE-2019-1110

XVII European Conference
ECSMGE 2019
Reykjavik
on Soil Mechanics
and Geotechnical Engineering

Risk assessment and dams – Recent developments and applications

Approche fiabiliste et barrages – Développements récents et applications

S. Lacasse
Norwegian Geotechnical Institute (NGI), Oslo, Norway

F. Nadim, Z.Q. Liu, U.K. Eidsvig, T.M.H. Le, C.G. Lin[1]
Norwegian Geotechnical Institute (NGI), Oslo, Norway
[1]NGI and Tongji University, Shanghai, China

ABSTRACT: Increasingly, society and standards require "risk-informed" decisions. The paper demonstrates the benefits of implementing reliability and risk concepts in dam engineering as a complement to conventional deterministic analyses. Reliability evaluations can range from qualitative estimates, to simple statistical evaluations, to full probabilistic modelling of the hazards and consequences for a system of dams. The paper gives an overview of basic concepts of reliability-based approaches and illustrates their use with three case studies. The paper discusses the strengths of reliability-based analyses and key issues such as tolerable and acceptable risk, the meaning of factor of safety, the targets for a margin of safety and the selection of characteristic value for analysis. Reliability-based approaches provide useful insight and complementary information. They enable the analysis of complex uncertainties in a systematic and more complete manner than deterministic analyses alone, both for the design of dams and for their safety evaluation during their lifetime. Reliability and risk-based approaches assist with preparing engineering recommendations and making decisions.

RÉSUMÉ: La société et les normes exigent de plus en plus de décisions "fondées sur le risque". L'article démontre les avantages de la mise en œuvre de concepts de fiabilité et du risque dans le dimensionnement de barrages comme un complément aux analyses déterministes classiques. Les évaluations de fiabilité peuvent être des estimations qualitatives, de simples évaluations statistiques, ou une modélisation probabiliste complète des aléas et des conséquences pour un système de barrages. L'article donne un aperçu des concepts de base des analyses fiabilistes et illustre leur utilisation avec des études de cas de trois barrages en remblai rocheux. Le document discute des points forts de l'approche fiabiliste et les questions clés telles que le risque tolérable et acceptable, la signification du facteur de sécurité, les objectifs pour une marge de sécurité et le choix de la valeur caractéristique pour l'analyse. L'article conclut que les approches basées sur la fiabilité fournissent des informations complémentaires utiles et permettent d'analyser des incertitudes complexes de manière systématique et plus complète que les analyses déterministes, tant pour la conception des barrages que pour leur évaluation pendant leur durée de vie. Des approches basées sur la fiabilité et le risque aideront à préparer les recommandations d'ingénierie et à prendre de bonnes décisions.

Keywords: Risk assessment, risk management, dams, reliability index, failure probability

1 INTRODUCTION

Karl Terzaghi (1929; 1961) wrote: *"Soil engineering projects require a vast amount of effort and labor securing only roughly approximate values for the physical constants that appear in the equations. The results of the computations are not more than working hypotheses, subject to confirmation or modification during construction. In the past, only two methods have been used for coping with the inevitable uncertainties: either adopt an excessively conservative factor of safety, or make assumptions in accordance with general, average experience. The first method is wasteful; the second is dangerous."*

Concepts of reliability and risk analyses are presented to illustrate the benefits of using reliability-based concepts for taking into account uncertainties and for the design and follow-up of dams. The paper discusses the significance of level of safety and the influence of uncertainties on the computed factor of safety. The need to use reliability-based approaches has risen because society and standards require more than before "risk-informed" design and "risk-informed" decision-making (ISO2394:2015). Reliability evaluations can range from qualitative estimates and simple statistical evaluations to full probabilistic modelling of the hazards and consequences for a system of dams.

The paper illustrates the use of the reliability-based approach with three case studies of embankment dams and summarizes the lessons learned. Key issues such as tolerable and acceptable risk and targets for a margin of safety are discussed. Reliability-based approaches provide useful insight and complementary information, and enable the analysis of complex uncertainties in a more systematic manner than deterministic analyses alone, both for the design of dams and for the safety evaluation during their lifetime. Reliability and risk-based approaches assist with the preparation of engineering recommendations and with making decisions.

2 A PRACTITIONER'S APPROACH TO UNCERTAINTY

During a design, the engineer always looks at the safety to know whether or not the foundation or the geotechnical structure can fail to perform adequately under the applied loads. The many uncertainties affecting geotechnical calculations need to consider their effect on the performance.

Silva *et al* (2008) combined historical and subjective probabilities (Fig. 1) to establish an approximate correlation between safety factor and failure probability. The diagram was devised for engineering practice. The figure is updated from Lambe (1985) and Baecher and Christian (2003), and compiles data from over 75 projects spanning over 4 decades. The projects include zoned and homogeneous earth dams, tailings dams, natural and cut slopes and earth retaining structures. The annual probabilities of failure for the different case studies were quantified iteratively, through experience, engineering judgment and published (historical) statistics.

Figure 1 cannot be used to either establish annual probabilities of failure in a design or verification situation or a relationship between factor of safety and probability of failure. It illustrates clearly, however, the effect of the uncertainties on the perceived factor of safety. Silva *et al* (2008) classify the structures into categories based on a judgment of the level of engineering. The level of engineering was established subjectively on the basis of the design practices (investigation, testing and analysis), documentation, construction, operation and monitoring:

Category I:
Facilities designed, built, and operated with state-of-the-practice engineering;
Category II:
Facilities designed, built, and operated using standard engineering practice;
Category III:
Facilities without site-specific design and substandard construction/operation;
Category IV:
Facilities with little or no engineering.

The family of curves were anchored on two sets of coordinates: a factor of safety of 1.5 for an annual probability of failure of 10^{-4} based on the historical performance of earth dams designed and constructed with conservative engineering practice (Baecher *et al* 1980; Whitman 1984; Christian *et al* 1992); and a 50% annual probability of failure for a safety factor of unity, based on a normally distributed uncertainty in the factor of safety (Vick 1994).

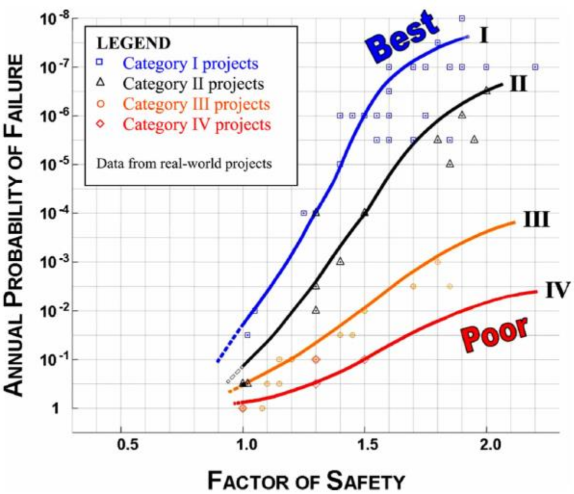

Figure 1. Practitioners' view of how factor of safety varies with uncertainty (Silva et al 2008)

Figure 1 suggests that, in the view of three practitioners (Silva *et al 2008*), a factor of safety of 1.5 can have an annual failure probability between 10^{-7} and 10^{-2} depending on the uncertainties, and a factor of safety of 1.3 an annual failure probability between 10^{-4} and 50%. The range of perceived failure probability for a "given" factor of safety is extremely wide.

3 UNCERTAINTIES AND EVENT PROBABILITIES

A statistical distribution is a practical tool to quantify uncertainty (Fig. 2), with a mean μ, a standard deviation (*SD*) and a coefficient of variation, *CoV*, which is an expression of the size of the standard deviation with respect to the mean ($CoV = SD/\mu$).

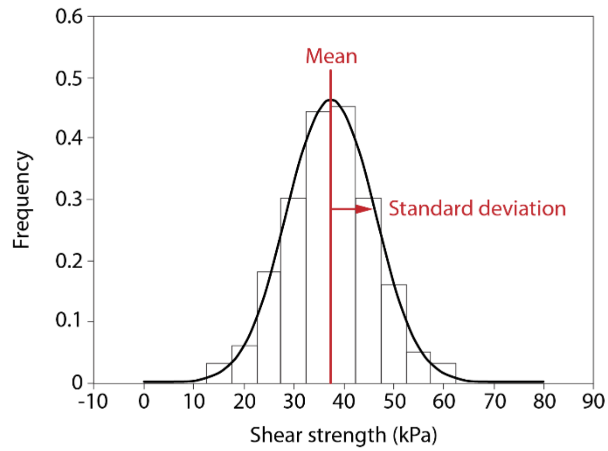

Figure 2. Uncertainty in a parameter

Uncertainties are either aleatoric or epistemic. *Aleatoric* uncertainty (also known as variability) is the natural randomness of a property or a load, e.g. soil strength or rainfall. The aleatoric uncertainty cannot be reduced. *Epistemic* uncertainty is the uncertainty due to lack of knowledge, e.g. measurement and method uncertainty. The epistemic uncertainty can be reduced by, e.g. increasing the number of tests or measurements, improving the measurement method and/or verifying the calculation procedure with model tests. Since the uncertainty is never zero, there is always a finite, even if small, probability that a failure may occur.

Both load and resistance have uncertainties (Fig. 3). Failure probability relates to the overlap of the two uncertainty distributions. Høeg and Murarka (1974) illustrated how uncertainties influence the probabilistic design of a retaining wall based on partial safety factors.

Uncertainties, their sources and their treatment could be a paper in itself. The quantification of uncertainties is not part of this paper. The reader can refer to several books and papers on this subject, including Ang and Tang (2007), Baecher and Christian (2003), Keaveny *et al* (1990), Lacasse and Nadim (1996), Lacasse *et al*

(2017), Nadim (2015), Tang (1973; 1984; 1987) and Uzielli *et al* (2006). This paper concentrates on risk-based approaches and the insight they bring to improve the design and safety of dams. The use of risk-based approaches is illustrated with case studies of dams.

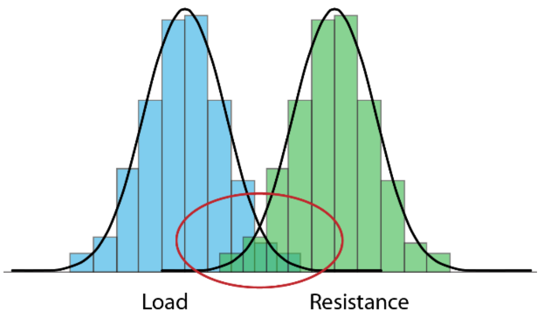

Figure 3. Potential overlap of load and resistance

The engineering literature (*e.g.*, Morgenstern 1995; Vick 1994) identifies three ways of estimating annual event probabilities: (1) based on the frequency calculated from observations (historical data); (2) derived from probability theory (reliability-based design with some mathematical modelling); and (3) using and quantifying, where possible, expert judgment (subjective probabilities).

Benjamin and Cornell (1970) stated that *"the sources of the probability [estimates] may include observed frequencies, deductions from mathematical models, and in addition, measures of an engineer's subjective degree of belief regarding the possible states of nature"*.

4 SAFETY FACTOR

Figure 4 shows that a design with a high factor of safety can have higher annual failure probability than another with a lower factor of safety. A higher safety factor, as commonly calculated, does not necessarily imply a smaller risk, because it is affected by the uncertainties in the analysis. The curves in Figure 1 reflect the concept in Figure 4: a design with a high factor of safety can have a higher probability of failure than another with a lower factor of safety. A higher factor of safety does not imply a smaller risk. Duncan (2000) pointed out that: *"Through regulation or tradition, the same value of safety factor is applied to conditions that involve widely varying degrees of uncertainty. This is not logical"*.

Safety factor is therefore not a sufficient indicator of safety because it does not account for the uncertainties in the analysis.

Soil properties and quality of engineering are not the only sources of uncertainty. The methods used to calculate stability, displacements or bearing capacity have themselves significant uncertainties. Nadim and Lacasse (1992) gave an example where stability analyses were done with the so-called effective stress and total stress approaches on a 'contractant' and a 'dilatant' soil. The computed failure probability differed significantly for each approach, although the computed factors of safety for the dilatant material were nearly the same. The differences in the calculated safety factors and the nominal failure probabilities were due to the different uncertainties in soil parameters and calculation methods.

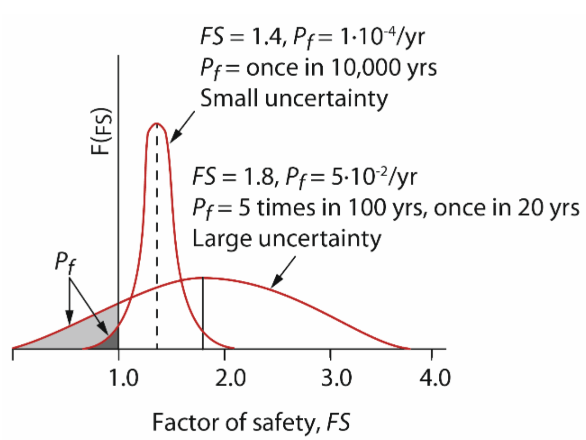

Figure 4. Safety factor and failure probability (P_f) of a slope (Lacasse and Nadim 1996)

5 RELIABILITY CONCEPTS

The terminology in this paper is consistent with the recommendations of ISO 31000:2018:

Danger (Threat):
Phenomenon that could lead to damage.
Hazard:
Probability that a danger (threat) occurs within a given period of time.
Exposure:
The circumstances of being exposed to a threat.
Vulnerability:
The degree of loss to a given element or set of elements affected by a hazard.
Risk:
Measure of probability and severity of an effect to life, health, property or environment.

Risk is the product of the probability of an event occurring (hazard) with the consequences due to this event. The consequences depend on exposure and vulnerability of the elements at risk. ISO 31000:2018 defines risk as the "*effect of uncertainty on objectives*". Consequences can be positive or negative and risk can be positively or negatively affected by changing circumstances.

5.1 Deterministic and probabilistic analysis

The terms "deterministic" analysis and "probabilistic" analysis are used.

- A deterministic system is one in which no randomness is involved in the estimate of future states of the system. A deterministic analysis aims at demonstrating that a facility is tolerant to identified faults or hazards within a "design basis", and evaluates a "nominal" performance. The approach does not consider the full range of possible outcomes nor quantify the likelihood of each of the outcomes. Deterministic scenario(s) may underestimate the risk.
- A probabilistic analysis aims at providing an estimate of the risk associated with a facility, and an estimate of the uncertainties involved. Probabilistic risk assessments help understand and account for the uncertainties. Discussing the uncertainties will, in any case, usually promote a debate that should lead to more insight and robust decisions.

While a deterministic analysis considers the impact of a single scenario (and a single set of input data), a probabilistic analysis attempts to include all possible scenarios, their likelihood and impact. A probabilistic analysis is comparable to series of sensitivity analyses (many thousands, even millions, of analyses).

5.2 Margin of safety

The objective of a safety assessment is to demonstrate that the risk associated with a facility is acceptable. The conventional way is to use a "deterministic" safety factor, FS. A safety factor of 1.5, for example, is often used to account for the combination of uncertainties in the ground, in the analysis parameters and the calculation method. There is a general perception that a design with a safety factor FS ≥ 1.5 has to be "safe".

Reality is not so simple. A safety factor 1.5 represents a spectrum of failure probabilities, which depend on the uncertainties in the analysis.

In a safety assessment, the engineer aims to quantify the margin of safety (M). Margin of safety is defined as:

$$M = Resistance - Load \tag{1}$$

With $M > 0$, the structure is safe; if $M \leq 0$, the structure is not safe. There is also an uncertainty in the safety margin (Fig. 5), and the failure probability, P_f, is the zone under the probability distribution of M where $M \leq 0$.

5.3 Reliability-based design

There are three approaches to design:

- The "Working stress" design (WSD) is the traditional approach based on an overall factor of safety, and has been used for a long time.
- Modern design codes are based on partial safety factors (or coefficients): the LRFD

approach (Load and Resistance Factor Design) in North America and the characteristic values and "partial safety factors" approach in Europe. The partial safety factors (also in the LRFD approach) are used to reflect the level of uncertainty and/or the relative importance of a particular parameter in design.

– Reliability-based design (RBD) using a target annual failure probability or target reliability index to verify margin of safety.

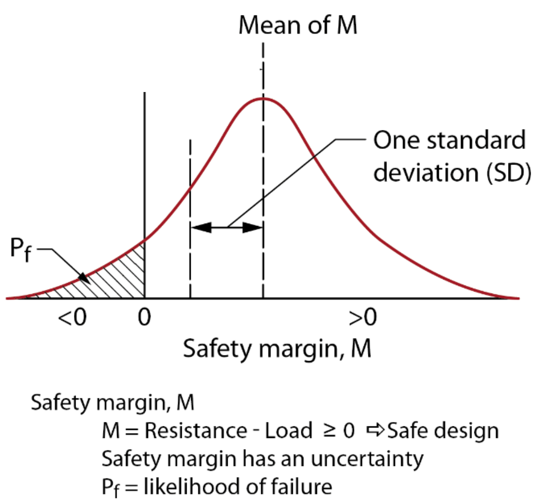

Figure 5. Safety margin and failure probability

5.3.1 Reliability index and failure probability

An alternative to using failure probability is to express the safety target in terms of an annual reliability index, β. Reliability index has a more positive connotation than failure probability and the two terms are directly correlated. Reliability index refers to the number of standard deviations between the mean safety margin and failure (M=0) in Figure 5. Reliability index is defined as:

$$\beta = \frac{M_{mean}}{SD} \tag{2}$$

Figure 6 gives the relationship between failure probability and reliability index for a normally distributed safety margin. For example, a reliability index (β-value) of 3.7 corresponds to a failure probability (P_f) of 10^{-4} and a β-value of 4.3 to a P_f of 10^{-5}. Similar curves exit for other distributions, e.g. the lognormal and triangular distributions.

Practice should use the concept of reliability index. Failure probability reminds of failure, reliability index refers to the reliability of a dam ("*Fiabilité*" in French, which is a synonym for "trustworthiness").

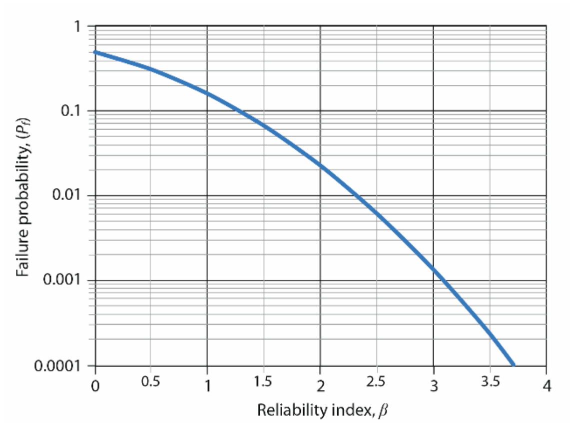

Figure 6. Relationship between probability of failure, P_f, and reliability index β (normal distribution)

5.3.2 Deterministic or probabilistic safety target

For large dams, there are often discussions of the safety target to achieve, and whether the safety target should remain the same during the entire life. The individuals downstream of a dam should not be exposed to a higher risk with time, and any potential environmental damage should not increase with time.

What should be the safety target during the life of a dam? Is a fixed deterministic safety factor sufficient to ensure the same safety margin throughout the lifetime of a dam? A safety level can be reflected in a constant annual failure probability, but not necessarily in a constant safety factor, because the uncertainties are not the same during the life of a dam, the likelihood of events may change and the consequences change with the downstream development.

A target annual failure probability, on the other hand, allows a more consistent comparison of the safety margin at different times of the life of a dam. A dam in operation for 50 years, represents 50 years of evaluated experience, not unlike a prototype test on site for 50 years under operation and environmental loads. In most cases, the uncertainties under design and construction will have reduced with time as more information and data become available, and as the performance of the dam is experienced over 50 years. The reliability-based approach can account for the observations and experiences during the course of operation of the dam.

5.4 Risk assessment and management

Risk management is the process of identifying, analysing and assessing risks to enable informed decisions on accepting or treating and controlling risks to minimize them. Risk management integrates the recognition and assessment of risk with the development of appropriate risk mitigation strategies.

Risk management comprises six main tasks: (a) Danger or hazard identification; (b) Causal analysis of the dangers or hazards; (c) Consequence analysis, including vulnerability analysis; (d) Risk assessment combining hazard, consequence and uncertainty assessments; (e) Risk evaluation of whether the risk is acceptable or not; and (f) Risk treatment (or risk mitigation).

Risk management has been formalised into a framework by ISO 31000:2018, with an integrated process involving communication and consultation on the one hand, and monitoring and review on the other hand. The process systemizes knowledge and uncertainties, to evaluate the significance of risk and for comparing options. In 2018, ISO added a "recording and reporting" requirement, and the entire risk assessment and management process is assimilated to a revolving circle.

There are several methods to do risk assessment, from simple qualitative risk matrices to more advanced numerical tools. Lacasse and Nadim (2007) summarized many of the methods in detail, and the methods are only briefly mentioned herein.

5.4.1 Qualitative methods

The most common tool is the "traffic-light" matrix (Fig. 7). Such qualitative matrices are very useful, especially when assessed through the consensus of several individuals with different expertise. Over the years, the 5x5 matrix has become more popular than the original 3x3 matrix. In the matrix, green designates "Low risk", red "High risk" and orange a situation in between, "Medium risk". Such qualitative estimates are useful, even recommended, as a first pass tool to establish whether or not a more detailed analysis is needed (if scenarios fall in the red or orange zones).

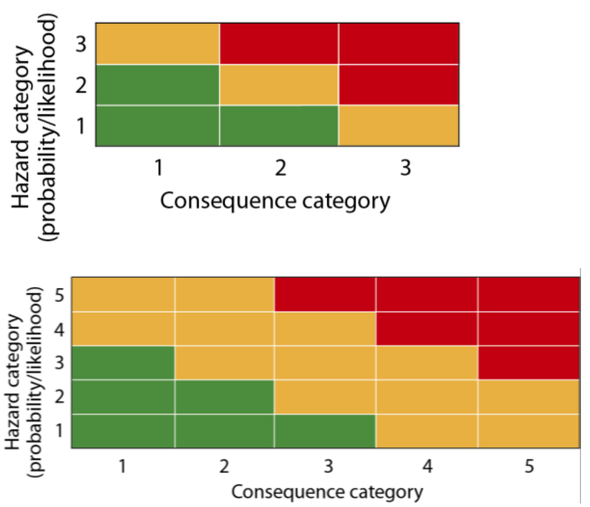

Figure 7. Qualitative 3 × 3 and 5 × 5 risk matrix: Hazard categories 1 to 3 or 1 to 5 (Very Low to Very High hazards); Consequence categories 1 to 3 or 1 to 5 (Very Small to Very Large consequences)

Such matrices can be implemented in e.g. a macro-operated Excel sheet (Langford *et al* 2019). It is important in such qualitative evaluation to define and use unambiguously and consistently the definitions of 'very low', 'low', 'medium', 'high' and 'very high' hazards and of 'very small', 'small', 'medium', 'large' and 'very large' consequence (or impact). Other qualitative methods, devised mainly for large dams and other critical facilities, include the Life Cycle Inventory and Life Cycle Assessment, also known as "Cradle to Grave" analysis (US EPA 2010) and the "Failure Modes and Effects Analysis" (FMEA) and the Failure Modes, Effects and Criticality Analysis (FMECA) (USACE 2011; USACE 2014; FEMA 2015).

5.4.2 Quantitative methods

In a quantitative probabilistic analysis, the same equation for load and resistance is used as for the deterministic calculation. The difference is that the material and load properties are described by a probability distribution function with a mean and standard deviation, that an additional variable is introduced, the method uncertainty.

Quantitative methods include: Event tree analysis, Fault tree analysis, Bayesian updating and the First Order Second Moment (FOSM) method. More complex tools are: Monte-Carlo simulations, Bayesian networks, the First and Second Order Reliability Method (FORM and SORM) and system reliability analysis, such as SYSREL Schneider (1997) and Lacasse and Nadim (2007) summarized many of the methods in detail. One can also combine two or several approaches to obtain reliability estimates, e.g. a probabilistic analysis of slope stability with an event tree analysis covering all plausible breach scenarios.

5.5 Acceptable and tolerable risk

Risk acceptance criteria are difficult to set, for both deterministic and probabilistic analyses. Acceptable risk refers to the level of risk requiring no further reduction, and is the level of risk society desires to achieve. Tolerable risk refers to the risk level reached by compromise in order to gain certain benefits. A construction with tolerable risk requires no action or expenditure for risk reduction, but it is desirable to control and reduce the risk if the economic and/or technological means for doing so are available.

A Frequency-Consequence chart (*F-N* chart) is a practical way to present risk level and compare different facilities. The *F-N* curves relate the annual probability *F* of an event to the number of fatalities *N*. The term "*N*" can be replaced by other measures of consequences such as costs. Figure 8 presents the Whitman (1984) *F-N* chart.

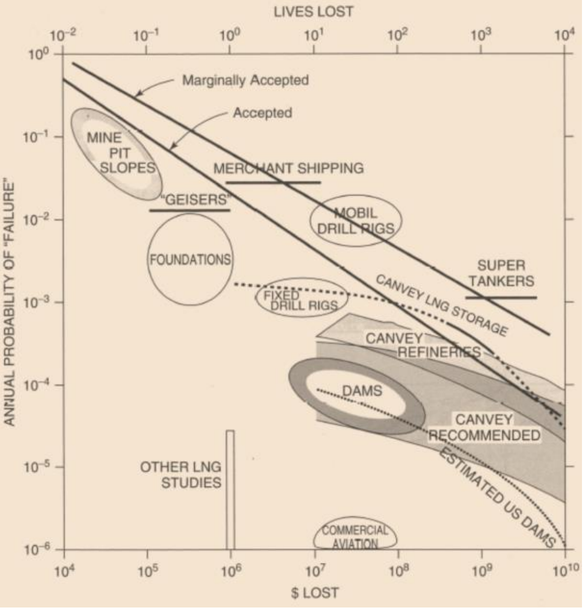

Figure 8. Whitman (1984) curves of acceptable and marginally acceptable risk (adapted by Baecher and Christian 2003) (note: 1984 US dollars).

Guidelines have been suggested by several countries. Some are for dams, some for man-made slopes, some more general (Fig. 9). Although there are differences, the annual acceptable risk level centres around 10^{-4} for ten fatalities. Figure 10 illustrates the range of published guidelines and the Hong Kong criterion for man-made slopes, which is one of the most frequently used criteria. The F-N diagram in Figure 9 is more stringent than Whitman's. The area to the right where the number of fatalities is greater than

1000 requires detailed assessment and reflects risk aversion in cases of very high number of fatalities.

The demarcation between acceptable and unacceptable risk is usually a gradual transition (Fig. 11). In Figure 11, the zone for *F* between 10^{-3} and 10^{-4} and 1 to 8 mortalities seems to belong to two categories. If a risk estimate should fall in that zone, the most severe action (red line) should be applied.

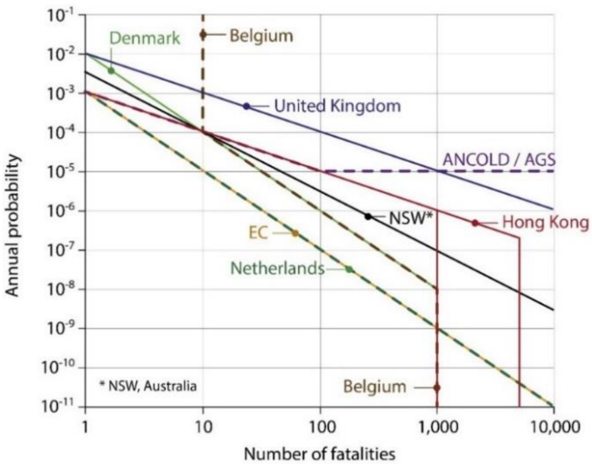

Figure 9. F-N risk guidelines in different countries (K. Ho Pers. comm. Hong Kong Nov. 2008)

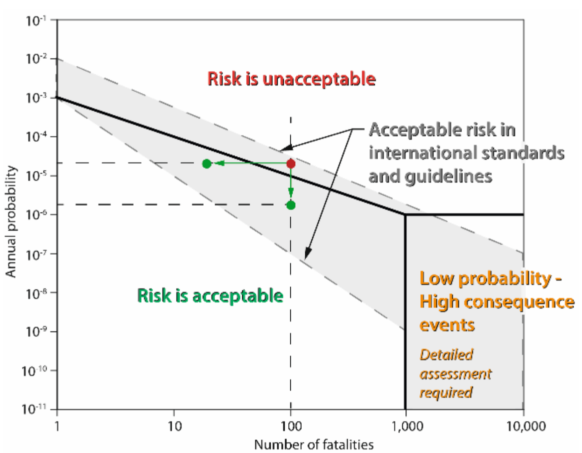

Figure 10. Range of risk guidelines (green dots illustrate how to reduce risk)

Hypothetical reductions of the failure probability for the same consequence of the consequences or for the same failure probability are illustrated with circles in Figure 10. It is not possible to show such evolution with a fixed factor of safety used as safety target. One can also define an *ALARP* zone on the F-N chart, where the risk level is to be kept "As Low As Reasonably Practicable". The *ALARP* zone describes a level of risk that cannot be reduced further without efforts and cost being disproportionate to the benefit gained or where the solution is impractical to implement.

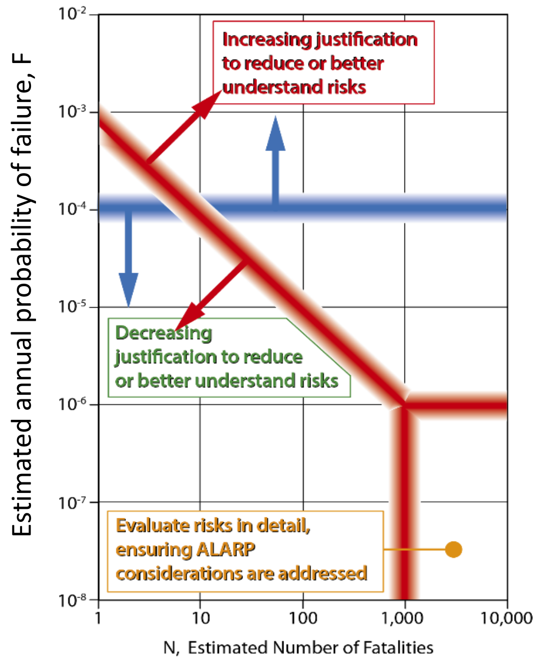

Figure 11. US Bureau of Reclamation 2011 guidelines

An annual failure probability of 10^{-4} has an actual significance. Figure 12 shows the mortality rate in Canada, due to all causes, as a function of age (https://www150.statcan.gc.ca/). At age 5-10, the probability of dying in the next year is 1/10,000 or 10^{-4}. At age 40, the probability increases to 1‰, at age 65 to 1%. As we reach a respectable age of 90, the probability of dying in the next year is close to 10%.

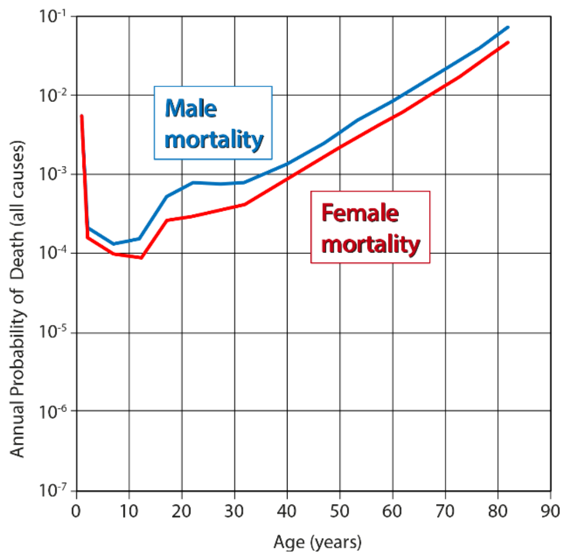

Figure 12. Annual probability of dying (Statistics Canada)

6 CASE STUDY: EMBANKMENT DAM UNDER WINTER CONDITIONS

6.1 Dravladalen Dam

The 340-m long rockfill embankment Dravladalen Dam has a height of 29 m and a reservoir of $58 \cdot 10^6$ m^3. The dam, built in 1971−1972, is founded on rock and has a till core. The dam is designed for a 1,000-year flood. Leakage was observed from the early stages of impoundment, but only small deformations were recorded. The "normal" seepage through the till core, based on laboratory seepage measurements, was 3 to 8 l/s. In 1994, the recorded leakage was 11 to 13 l/s. In 2016, the leakage under full reservoir was on average 5 to 6 l/s, and the leakage water was clear (no discernible fines). The dam is classified in the highest consequence class in Norway.

6.2 Reliability approach

The probabilistic analyses were carried out with two approaches: (1) event tree analysis, and (2) Bayesian network[1] combined with Monte Carlo simulations. The techniques are described in e.g. Hartford and Baecher (2004), Baecher and Christian (2003) and Lacasse *et al* (2017). Each of the methods uses nine steps (modified after Vick 2002; Høeg 1996):

1. Review of field performance and history.
2. Dam site inspection and data review.
3. Failure mode screening.
4. Agreement on descriptors of uncertainty.
5. Event tree construction.
6. Probability estimate at each node of the tree.
7. Calculation of annual probability of breach.
8. Evaluation of results.
9. Iteration and documentation.

The analyses were done in a workshop mode by bringing together 18 "experts" with knowledge about the dam, the hazards and risks involved, the dam construction and dam behaviour in general. The participants included dam owners, engineers responsible for the dam operation, hydrologists, earthquake specialists, reliability specialists, consultants and regulatory bodies. One person in charge of the day-to-day follow-up of the dam was at the dam site and available to answer questions. The format of a workshop was very useful to assess and discuss the estimates of probability. The probability estimates for the event tree and Bayesian analyses were set with the help of (1) statistical estimates based on past observations (actual data); (2) engineering models based on physical processes, e.g. stability analyses (including parameter uncertainties); and (3) expert judgment based on knowledge and evaluated experience. Vick (2002) suggested that: *"The collective judgment of experts, structured within a process of debate,*

[1] Bayesian network is an emerging method for reasoning and modelling under conditions of uncertainty. The method has been applied to avalanche risk, design of early warning system for landslide hazard mitigation, rock slope failure, dam risk analysis, earthquake risk management and multi-hazard, multi-risk assessment. Liu *et al* (2015) presented a summary of the uses so far.

can yield as good an assessment of probabilities as mathematical analyses".

The process of hazard and risk assessment for dams in a workshop format, where event trees are constructed for different plausible failure scenarios, helps identify the weak points in a complex system and/or in a reasoning and enables one to make the system more robust through, for example, hypothetical remediation measures for discussion sake. The probabilities assigned to each node in the event trees was extensively discussed in plenum at the workshop initially. There was often disagreement in the values assigned. Consensus was reached through argumentation, verification of hypotheses, consultation of additional information and discussion. The event tree analyses went through two or three iterations before the event probabilities on the tree branches and the final failure probabilities were determined.

6.3 Failure mode screening

One of the essential parts of a review of the possible failure modes before the construction of event trees. After discussion, the following mechanisms and triggers were examined:

Weaknesses in the dam or dam system:

- Internal erosion
- Slides in upstream and downstream slope
- Rockslide in reservoir causing overtopping
- Plane of weakness in bedrock foundation
- Operator error

External triggers:

- Flood
- Extreme snow/ice in the winter
- Earthquake
- Melting of glacier causing flood in reservoir
- Sabotage/terror
- Meteors or plane crashing into the dam

Before the workshop, NGI (2004) had checked that the safety was more than adequate against a rock slide triggering a tsunami in the Dravladalen Dam reservoir. The stability of the rock foundation was also checked for the high quality gneiss and granite rock foundation, with foliation dips upstream and no weakness planes downstream. Meteors and plane crash at the location of Dravladalen Dam were estimated to have occurrence probability of less than 10^{-7}/ year.

6.4 Results of analyses

Tables 1 and 2 present the results of two series of analyses, the first in 1996, the second in 2016, in terms of the annual probability of failure, $P_{f\,annual}$. The probabilistic analyses in 1996 led to the identification of an unforeseen mode of failure, which turned out to be the most critical failure mode (Johansen *et al* 1996). Remediation measures were completed during the subsequent years. The rehabilitation included: new toe for the dam to increase drainage capacity; new slope protection downstream, with gentler slope; new dam crest; new concrete shelter for the approach channel to the spillway tunnel; new leakage monitoring system; and instrumentation of upstream slope and dam crest.

The 2016 analyses looked at each of the failure modes and at the effectiveness of the remediation measures in the period 1996-2016. Table 3 compares the event tree reliability results from the analyses in 1996 and 2016. For1996, the results of the two iterations are shown. In 2016, three iterations were done, and each gave approximately the same results. The Bayesian network analyses, combined with over 500 Monte Carlo simulations, gave essentially the same 'mean' annual probability of failure as the event tree analyses in 2016. The Bayesian network analyses provided in addition the distribution of the failure probabilities, with a mean value of P_f, maximum value and minimum value of P_f, as illustrated in Figure 13 for the scenario of 'ice and hard-packed snow blocking the spillway tunnel'.

Table 1. Results of event tree analyses of Dravladalen Dam in 1996

Scenario	Most probable failure mechanism	$P_{f\,annual}$ (it. 1)	$P_{f\,annual}$ (it. 2)
Earthquake	Overtopping due to settlement of dam	$< 1.5 \cdot 10^{-6}$	$< 1.5 \cdot 10^{-6}$
Winter flood	Overtopping due to plugging of spillway tunnel	$3 \cdot 10^{-3}$	$4 \cdot 10^{-4}$
Int. erosion	Failure in downstream slope and toe *	$P_{f\,life} = 5 \cdot 10^{-4}$	$P_{f\,life} = 5 \cdot 10^{-5}$
Sabotage	Overtopping	$< 1 \cdot 10^{-5}$	$< 1 \cdot 10^{-5}$
All scenarios	$P_{f\,annual}$ without internal erosion	$< 3 \cdot 10^{-3}$	$< 4 \cdot 10^{-4}$

* In 1996, the lifetime probability of failure was calculated for the internal erosion case.

Table 2. Results of event tree analyses of Dravladalen Dam in 2016

Analysis		$P_{f\,annual}$ (last iteration)
Internal erosion		$4.7 \cdot 10^{-6}$
Flood	*Winter*: Ice and hard-packed snow blocking tunnel	$2.4 \cdot 10^{-7}$
	Summer: Glacier melting in reservoir	$5.4 \cdot 10^{-6}$
Earthquake		$9.0 \cdot 10^{-8}$
All geotechnical and natural hazards scenarios		$1.0 \cdot 10^{-5}$
Sabotage/terror		$2 \cdot 10^{-7}$

Table 3. Comparison of the results of event tree analyses in 1996 and 2016

Analysis		$P_{f\,annual}$ - 2016	$P_{f\,annual}$ - 1996
Internal erosion		$4.7 \cdot 10^{-6}$	$5 \cdot 10^{-5}$ (life)
Flood	*Winter:* Ice and hard-packed snow blocking tunnel	$2.4 \cdot 10^{-7}$	$4 \cdot 10^{-4}$
	Summer: Glacier melt in reservoir	$5.4 \cdot 10^{-6}$	--
Earthquake		$9.0 \cdot 10^{-8}$	$<1{,}5 \cdot 10^{-6}$
All geotechnical and natural hazards scenarios		$1.0 \cdot 10^{-5}$	$<4{,}2 \cdot 10^{-4}$
Sabotage/terror		$2 \cdot 10^{-7}$	$<1 \cdot 10^{-5}$

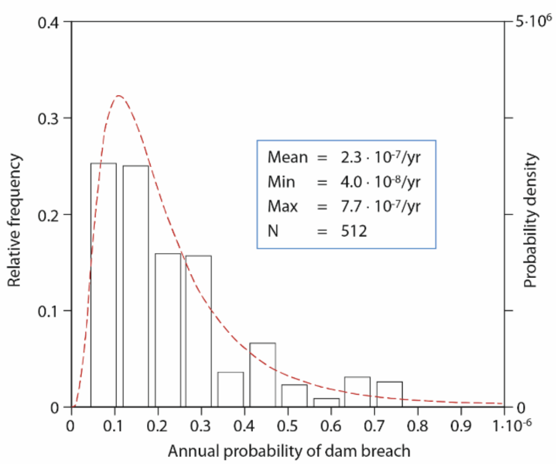

Figure 13. Dam Dravladalen: Annual failure probability from Bayesian Network and Monte Carlo analysis

Figure 13 gives the histogram of annual probabilities of failure and the best lognormal distribution fit, and gives the number (N) of Monte Carlo simulations done. The new risk assessment in 2016 showed that the annual failure probability for this failure mode was reduced by two to three orders of magnitude because of the implementation of the mitigation measures.

6.5 Summary

The application of reliability concepts can be useful for ensuring safe and cost-effective dam design and rehabilitation. The annual probability of failure for Dravladalen Dam in 2016 was estimated as 10^{-5} (once in 100,000 years). The annual probability of failure had been estimated as $0.4 \cdot 10^{-3}$ in 1996. The 1996 analyses identified a new failure scenario ('ice and hard-packed snow blocking spillway tunnel'), which had been overlooked in the deterministic design. The probabilistic analyses demonstrated the effectiveness of the mitigation measures implemented in the period 1996-2012.

An annual probability of failure of 10^{-5} is lower than the statistical annual probability of failure values reported by ICOLD and the published values for acceptable risk for dams. A consequence analysis, not reported herein, established that 200 to 300 persons could be affected by a dam breach, but no lives would be lost in the case of a dam breach, due to the long warning time for this dam built in a remote area.

When all the breach scenarios were examined together, the consensus was that the most probable scenario that could lead to a breach was the sabotage/terror scenario, because it included larger uncertainties than the other scenarios and because there are today no security measures on Dravladalen Dam. The situation is however not believed to be critical because the dam is located in a very remote area with very difficult access, summer and winter. A recommendation was made (1) to maintain the leakage and displacement observations as they provide useful information for future evaluation of risk and (2) to establish measures to limit and control access to the dam.

7 CASE STUDY: 40 YEARS OF SATISFACTORY PERFORMANCE

7.1 Nyhellervatn Dam

The 650-m long rockfill embankment Nyhellervatn Dam with till core is 82.5 m high and has a reservoir capacity of $450 \cdot 10^6$ m^3. The dam, built in the period 1975−1979, is founded on bedrock. The dam is designed for a 1,000-year flood. The main Nyhellervatn Dam is classified in the highest consequence class in Norway. The stability analysis of the downstream slope gave a deterministic safety factor less than the required 1.5 after 40 years of operation. Throughout these 40 years, the dam has behaved satisfactorily, with no unexpected leakage, pore pressure increase or displacements. Leakage is monitored continuously, and reported in real time.

7.2 Reliability approach

The reliability analyses used the same approaches as for Dravladalen Dam. The three iterations gave essentially the same results. In addition, Monte Carlo analyses with the "SLOPE/W" were used to verify the stability of the upstream and downstream slopes.

7.3 Failure mode screening

After discussion during the workshop, the following mechanisms and triggers were considered:

Weaknesses in the dam or dam system:
- Internal erosion
- Slides in upstream and/or downstream slope
- Leakage at rock foundation undermining the core
- Plane of weakness in bedrock foundation
- Rockslide in reservoir causing overtopping
- Operator error

External triggers:
- Flood
- Extreme snow/ice in the winter
- Earthquake
- Wave and ice loading upstream on rip-rap
- Melting of glacier causing flood in reservoir
- Sabotage/terror
- Meteors or plane crashing into the dam

7.4 Results of analyses

7.4.1 Failure probability

Table 4 presents the results of the probabilistic event tree analyses in terms of the annual probability of failure, $P_{f\,annual}$.

The Bayesian network analyses, combined with 1000 Monte Carlo simulations, gave the same 'mean' annual probability of failure as the event tree analyses. The maximum and minimum annual P_f (mean=$3.7 \cdot 10^{-7}$/year, minimum of

$1.0{\cdot}10^{-8}$/year and maximum of $1.2{\cdot}10^{-5}$/year respectively) were obtained for the scenario of 'Leakage through cracks in the rock foundation causing erosion of the core' (row 2 in Table 4). The Nyhellervatn histogram and distribution were narrower than that for Dravladalen, indicating significantly lower uncertainty in the probabilistic estimate.

Table 4. Annual failure probability for Nyhellervatn Dam

Scenario	$P_{f\ annual}$
Internal erosion	$1.5{\cdot}10^{-6}$
Leakage through cracks in rock foundation causing erosion of the core	$3.7{\cdot}10^{-7}$
Flood	$6.5{\cdot}10^{-9}$
Wave loading on upstream slope	$1.0{\cdot}10^{-8}$
Ice loading on upstream slope	$1.0{\cdot}10^{-8}$
Earthquake	$2.7{\cdot}10^{-7}$
All geotechnical and natural hazards scenarios	$2.2{\cdot}10^{-6}$

7.4.2 Stability of downstream slope

The deterministic analyses of the stability of the downstream slope suggested two potential critical slip surfaces, a very shallow slip surface A, and a deeper slip surface B (Fig. 14). The following values were used for the deterministic analysis in design (φ' is effective friction angle, c' is effective cohesion and γ is total unit weight) (carefully assessed "representative" values for the entire rockfill were used at the time):

Till core: $\varphi' = 33°$, $c' = 10$ kPa, $\gamma = 23$ kN/m^3
Rockfill: $\varphi' = 45°$, $c' = 0$ kPa, $\gamma = 20.5$ kN/m^3

There is, however, a large uncertainty in the friction angle of the rockfill, as illustrated in Figure 14. The strength of a rockfill depends on many factors, including the effective stress on the slip surface, compacted rockfill porosity and quality of the rockfill material. Figure 15 (upper) presents the available data on the shear strength of rockfill, based on Leps' data (1970) and additional experimental data (EBL 2003/NGI 2002). Figure 15 (lower) shows NGI's recommendation and the requirement in Norway for the design of rockfill dams, as imposed by Norwegian regulatory body NVE (The Norwegian Water Resources and Energy Directorate, www nve.no).

Probabilistic analyses were run to include the effect of the uncertainty in the friction angle of various types of rockfill. The probabilistic analyses used a lognormal distributed random variable for the friction angle of the rockfill, with a best estimate mean (not a careful assessment) and minimum and maximum values (thus a truncated lognormal distribution). The range of values used for slip surfaces A and B is shown with the green arrows in Figure 15. The values selected for the analyses were well within the recommendation by EBL/NGI.

Table 5 compares the results of the deterministic and probabilistic stability analyses of the downstream slope under stationary conditions (stability under rapid drawdown, flood and earthquake loading was also verified). The probabilistic analyses showed that, using a plausible range of frictions angles shown in Figure 15, the downstream slope safety was adequate, even though the deterministic analysis gave a safety factor somewhat less than 1.5.

Table 5. Analysis of stability of downstream slope under stationary conditions ($FS_{required} = 1.5$)

Slip surface	**Deterministic FS**	**Probabilistic FS_{mean}**	**Failure probability**
A	1.51	1.58	$P_f < 10^{-7}$
B	1.42	1.32	$P_f = 7{\cdot}10^{-7}$

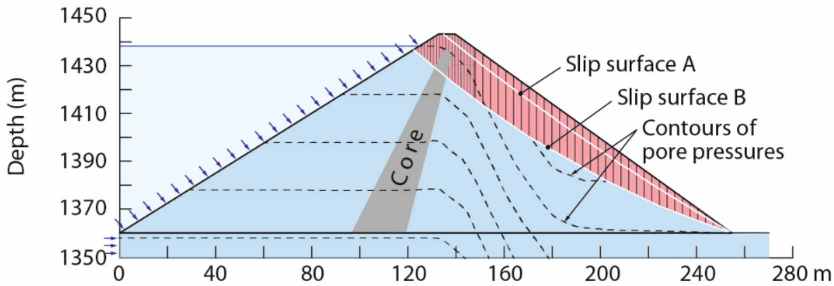

Figure 14. Stability analyses of downstream embankment, Nyhellervatn Dam

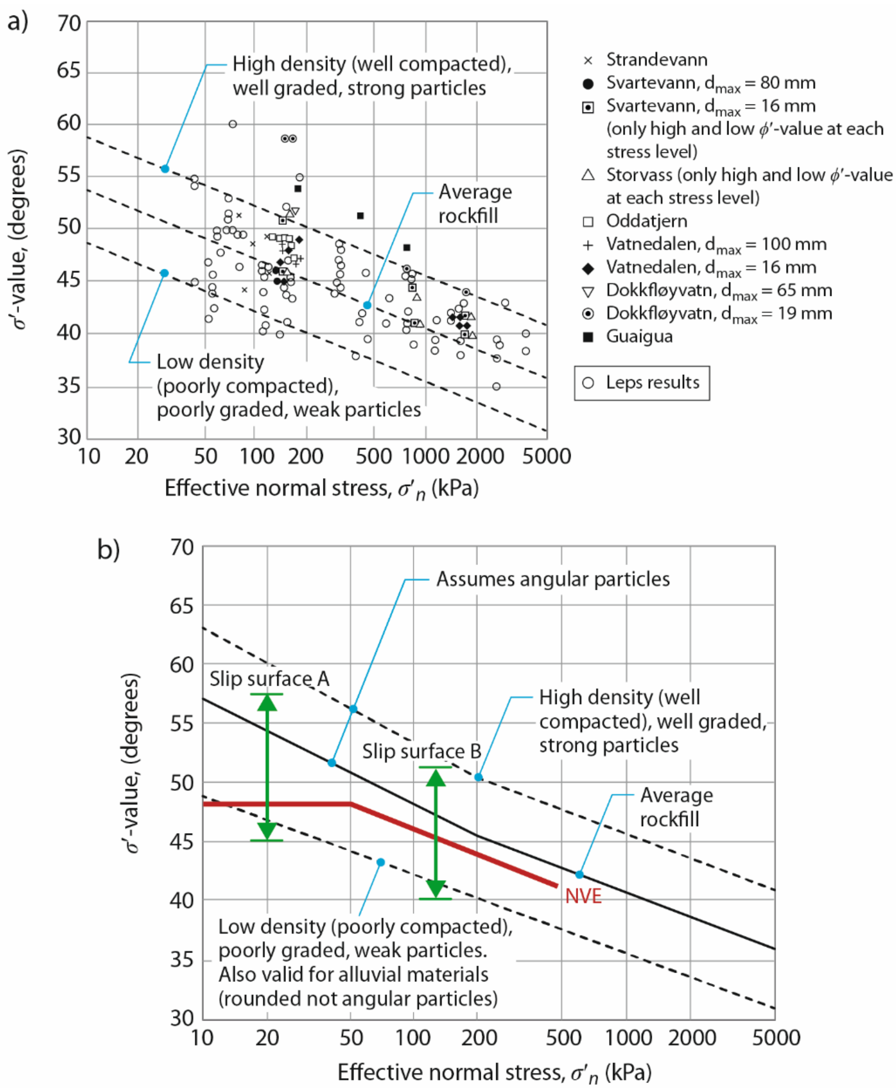

Figure 15. a) Data underlying recommendation in Lower diagram (EBL 2003/NGI 2002); b) Recommended friction angle values for rockfill materials (Høeg, K. Pers. comm. Bucuresti Inaugural Lecture 2008)

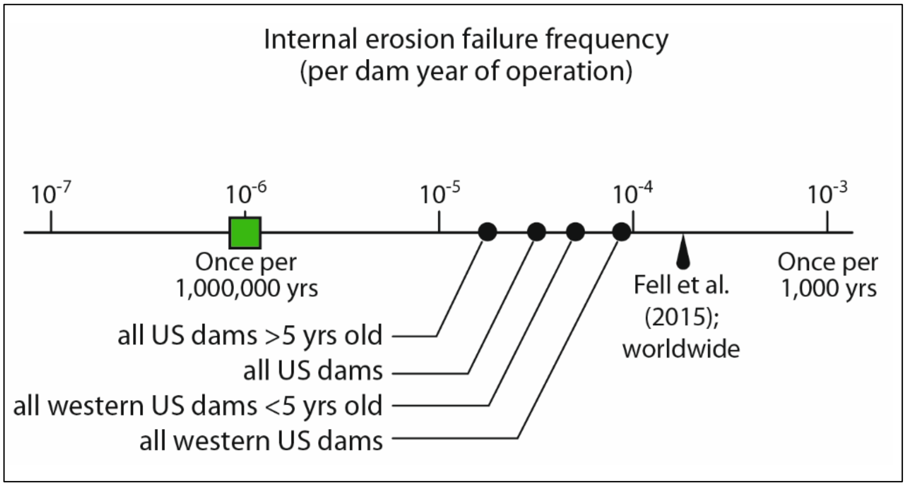

Figure 16. Comparison of failure frequency of Dam Nyhellervatn (green square) with statistics of breach due to internal erosion

7.4.3 Summary

The best estimate of the annual failure probability for Nyhellervatn Dam in 2017 was $2{\cdot}10^{-6}$ or lower, or a breach frequency of about once in one million years. The calculated failure probability is considerably lower than the breach frequency reported for dams in the literature, e.g. ICOLD's statistics. Figure 16 compares the failure probability with international statistics for dam breach due to internal erosion.

Nyhellervatn Dam is a solid, robust and safe dam. It is therefore important to pose the question of whether or not the dam should be rehabilitated to increase the safety factor of the downstream slope to meet the requirement of a constant safety factor, FS $\geq$ 1.5. The dam has been in operation for over 40 years, undergoing multiple drawdown and filling cycles, and experiencing very harsh as well as very mild summer and winters, and flooding events. Nyhellervatn Dam has not shown any signs of distress or unexpected behaviour. The 40 years of operation are in fact 40 years of a dam under full scale loading. The added knowledge should be taken into account in the deterministic analysis, but there is no mechanism to do this. Statistics for embankment dams (e.g. Fell *et al* 2015) also show that the majority of the failures occur in the first five operative years of the dam.

8 CASE STUDY: SYSTEM OF DAMS

8.1 The Nesjen Dams

The system of dams at Nesjen consists of one main rockfill embankment dam with till core, 50 m high, four secondary (saddle) dams (of much lower heights) and a separate spillway. The dams are briefly characterized in Table 6. All the dams are rockfill dams, except for Saddle Dam 1 which is a concrete buttress dam. The Main Nesjen Dam, built in the period 1966−1968, is founded on bedrock, and is designed for a 1,000-year flood. The Main Dam and Saddle Dams 2 and 3 are classified in the highest consequence class in Norway. The reliability analyses were done to evaluate the effects of different rehabilitation measures.

8.2 Reliability approach

The reliability analyses used for the Nesjen dams were the same approaches as for Dravladalen Dam. The three event tree iterations gave essentially the same results. In addition, first-order and second-order (FORM and SORM) probabilistic analyses were run to verify the stability of the concrete buttress Secondary Dam 1, but these

Table 6. Brief description of the dams at Nesjen

Characteristics	Main Dam	Saddle Dam 2	Saddle Dam 3	Saddle Dam 4	Saddle Dam 1
Dam length (m)	675	225	170	75	500
Volume (1000 m^3)	560	148	61	--	--
Maximum dam height (m)	50	19	15	15	10
Consequence class*	4	4	4	3	2

* Consequence class 4 is most severe consequence class in Norway

analyses are not reported herein. Failure probability of the concrete buttress was very low, and the consequence of a failure was much smaller than for the other dams (Consequence Class 2).

8.3 Failure mode screening

After discussion at the workshop, the following mechanisms and triggers were considered:

Weaknesses in the dam or dam system:
– Internal erosion
– Slides in upstream and/or downstream slope Leakage at rock foundation undermining the core
– Plane of weakness in bedrock foundation
– Rockslide in reservoir causing overtopping

Operator error *External triggers:*
– Flood
– Extreme snow/ice in the winter
– Earthquake
– Wave and ice loading upstream on rip-rap
– Melting of glacier causing flood in reservoir
– Sabotage/terror
– Meteors or plane crashing into the dam

8.4 Results of analyses

Table 7 and 8 present the results of the probabilistic analysis in terms of the annual failure probability, $P_{f\,annual}$, for the Main Dam before rehabilitation, and Table 8 the Main Dam after rehabilitation. Figure 17 compares the annual failure probability with the failure probability statistics for dam breach due to internal erosion. Even before rehabilitation, the Main Dam has a low failure probability. The different rehabilitation measures decreased importantly the annual probability of a breach.

For the Nesjen system of dams, it was important to not only analyse each dam separately but to also look at the risk associated with the dams in a system of one large dam and several smaller secondary dams. During the reliability analyses under extreme flooding, it was concluded that it was desirable to reduce the reservoir water level increase due to flooding at the Main Dam by allowing some damage due to overtopping of Saddle Dam 4. The consequences of an overtopping of Saddle Dam 4 are significantly smaller (no life loss) that the consequences of an overtopping of the Main Dam. There is also ample warning time downstream, with a planned overtopping of Saddle Dam 4.

Originally, the rehabilitation required that the core and the crest be raised equally for all dams. After the reliability analyses, the dam core and the dam crest of Saddle Dam 4 should be left at a lower elevation than the other 'Consequence Class 4' dams, such that the dam owner can plan for a controlled overtopping under an extreme flood event. On the basis of the reliability analyses, recommendations were made on the most effective rehabilitation measures for reducing the risk of internal erosion (Table 8), on warning systems in the case of a controlled overtopping of Saddle Dam 4, and the interrelationship between risk at the Main Dam and the raising of all secondary dams at the same level as that of the main dam.

Table 7. Annual failure probability for Nesjen Main Dam, before rehabilitation (last iteration)

Scenario	**Annual failure probability, $P_{f\ annual}$**
Internal erosion (iteration 1 and iteration 2))	$7.6 \cdot 10^{-5}$
Flood	$2.9 \cdot 10^{-7}$
Earthquake	$1.0 \cdot 10^{-8}$
Erosion in rock foundation	$5.0 \cdot 10^{-6}$
Total failure probability	$5.5 \cdot 10^{-5}$

Table 8. Annual failure probability for Nesjen Main Dam, after rehabilitation (last iteration)

Scenario	**Annual failure probability $P_{f\ annual}$**
Internal erosion (toe reinforcement, fibre cable in till and leakage monitoring)	$8.4 \cdot 10^{-6}$
Flood, raising dam crest and core	$1.8 \cdot 10^{-7}$
Flood, improving spillway	$2.0 \cdot 10^{-8}$
Earthquake	$1.0 \cdot 10^{-8}$
Erosion in rock foundation	$1.5 \cdot 10^{-6}$
Total failure probability	$9.1 \cdot 10^{-6}$

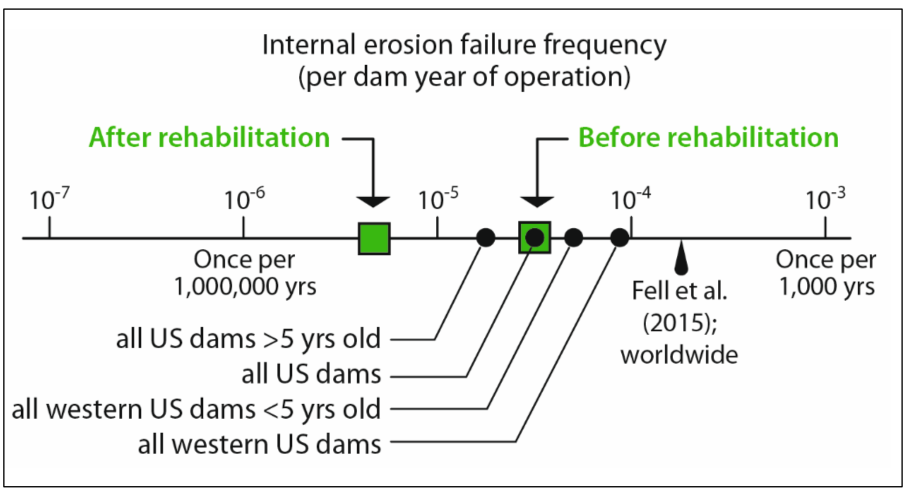

Figure 17. Comparison of failure frequency of the Main Nesjen Dam (two green squares) with statistics of breach due to internal erosion

8.5 Summary

The calculated best estimate of the annual failure probability for the Main Dam at Nesjen before the implementation of rehabilitation measures was $5 \cdot 10^{-5}$ (or about once per 50,000 years), and $1 \cdot 10^{-5}$ (or once per 100,000 years) after selected rehabilitation measures were implemented. The calculated probabilities were lower than the annual probabilities of dam breach compiled in international statistics for embankment dams (e.g. Fell *et al* 2015).

The reliability analyses shed light on the benefit of allowing overtopping of Saddle Dam 4 to reduce the failure probability of a breach of the Main Dam. In the case of an extreme flood situation, an overtopping of Saddle Dam 4 will reduce the reservoir water level at the Main Dam and immediately reduce the risk of a breach of the Main Dam. The consequences of a dam breach at the Saddle Dam 4 are significantly lower than for the Main Dam, and would not cause loss of human life. Also there is ample warning time for the people downstream.

Originally, the rehabilitation required that the core and the crest be raised equally for all dams. Rather, after the reliability analyses, the recommendation is that the dam core and the dam crest of Saddle Dam 4 should be left at a lower elevation than the other 'Consequence Class 4' dams at Nesjen, such that the dam owner can plan for a controlled overtopping of Saddle Dam 4, if necessary under an extreme flood event.

9 RECENT DEVELOPMENTS

9.1 Use of machine learning algorithms

Artificial intelligence techniques, like machine learning, can be used to predict soil behaviour.

The Three Gorges Dam reservoir is a landslide-prone area and the construction of the Three Gorges Dam dramatically increased landslide hazard in the area. A reliable early warning system would help reduce the risk associated with landslides. Such systems can be successful if one can forecast an imminent landslide.

Yang et al (2019a; b) proposed a novel machine learning model to predict landslide displacement in dam reservoirs, and applied the model to the Three Gorges reservoir. The machine learning model uses time series analysis and the Long Short Term Memory (LSTM)" neural network approach.

Long Short-Term Memory (LSTM) neural networks are a type of "Recurrent Neural Networks (RNN)" designed to model temporal sequences and time dependency more accurately than conventional RNNs. In simple words, LSTM has a memory block, which relates one time step to another. The memory block can retain or forget information. The 'Input gate' controls the flow of input activations into the memory cell. The 'Forget gate' controls whether the information from the previous time step is remembered or forgotten. During the process, the LSTM model learns rules from historical information and makes full use of this information. Yang *et al.* (2019a) described in more detail the architecture of the LSTM neural network used.

The displacement was decomposed into three components: a trend (T), a periodic (P) and system noise (N) components. The accumulated total displacement corresponds to $T + P + N$:

- The long term displacement (trend), controlled by 'internal' geological conditions such as lithology, geological structure and progressive weathering.
- The periodic short term displacement, influenced by two 'external' factors: rainfall and reservoir water level.
- The system error covering systematic errors during the deformation monitoring process.

The trend displacement was predicted using a cubic polynomial function.

The periodic displacement was predicted by a multivariate LSTM model based on the relationship among landslide displacement, rainfall and reservoir water level: seventy percent of the data were used to develop the model. To verify the performance of the new model, the latter 30% of the displacements were predicted by the model and compared with the measurements.

The performance of the LSTM model was validated with the observations of three typical "step-wise" colluvium landslides in the Three Gorges Dam Reservoir, and compared with other machine learning prediction models.

Figure 18 presents the results of the fitting of the model and the prediction if the periodic displacements for two of the validation landslides. Figure 19 presents the same comparison for the total accumulated prediction. The predicted values fitted well with the measured values during the training of the model and can then be used to predict the future behaviour. The model was able to reflect the dynamic evolution of landslide deformation by relating observations from one time step to the next, thus introducing a dynamic component in the analysis.

The application of the model to the three landslides demonstrates that the LSTM model gave a more reliable prediction of the observed landslide displacement than a static model. It was concluded that the new model can be used to effectively predict the displacement of colluvium landslides in the Three Gorges Reservoir area. Such reliable predictive models can become an essential component for implementing an early warning system and reducing landslide risk.

Overall, the proposed dynamic modelling approach, based on time series analysis and LSTM, can achieve a good prediction of displacements for slow and step-wise deformations. This dynamic method has the potential for broad application to predict landslide displacement in landslide-prone regions.

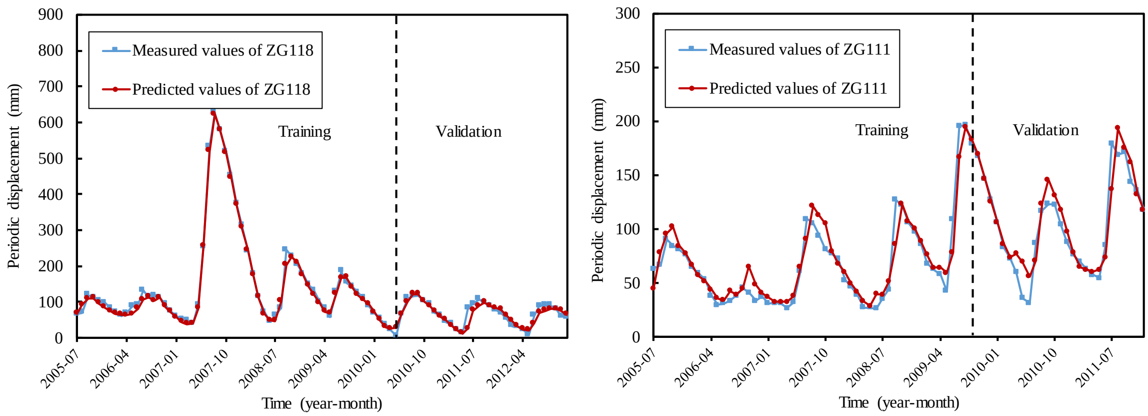

Figure 18. Periodic displacement for Baishuihe (left) and Bazimen (right) landslides (Yang et al 2019a; b)

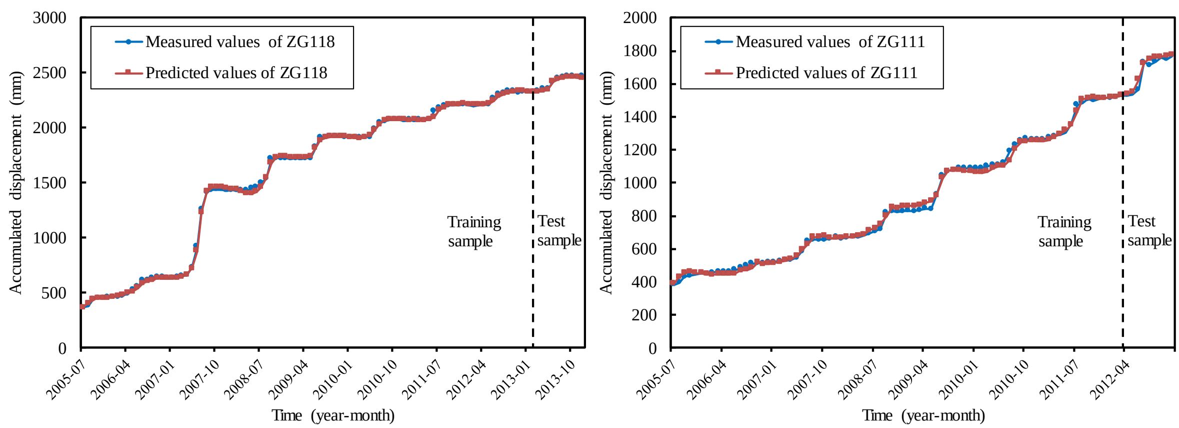

Figure 19. Total accumulated displacement for Baishuihe (left) and Bazimen (right) landslides (Yang et al 2019a: b)

9.2 The Observational Method

Terzaghi's and Peck's Observational Method should be used to greater extent for the design and the follow-up of dams on sites with complex geological conditions. The Observational Method (Peck 1969) includes several aspects of uncertainty and risk in geotechnical design, by looking at the mean and the uncertainty (*assessment of the most probable conditions and the most unfavourable conceivable deviations from these conditions*[2]), evaluating the hazards *(calculation of values of the same quantities under the most unfavourable conditions*) and preparing mitigation measures *(selection in advance of a course of action or modification of design for every foreseeable significant deviation of the observational findings from those predicted on the basis of the working hypothesis* and *modification of design to suit actual conditions*). Peck gave in 1996 the reason he published the Observational Method: *"My real interest, instead* [of theoretical research] *was in the ways our existing knowledge could be applied more effectively"*.

ICOLD (1993) strongly suggested that the Observational Method is desirable, even required, for seepage control and drainage treatment in a dam foundation. Information gained during foundation excavation and further investigations may significantly modify and improve the original design. The implementation of the Observational Method and stabilization measures helped reduce considerably the risk of instability for the Zelazny Most tailings dam in Poland on a geologically very complex foundation (Jamiolkowski 2014: Jamiolkowski *et al* 2008). The application of the Observational Method resulted in measures such as moving the dam crest upstream to flatten the average downstream slope, constructing stabilizing berms at the dam toe, and installing relief wells in the foundation to reduce pore water pressures. Other examples of the benefit of the monitoring and the Observational Method are presented in Table 9 for three Norwegian dams.

Table 9. Benefits of monitoring program for three dams in Norway (Lacasse and Höeg 2019)

Type of dam/Dam height/Year/Core	**Benefit of monitoring program**
<u>Moravatn Dam</u> Rockfill dam 77-m/1968 Moraine core	Confirmed need for rehabilitation from the high pore pressure in the dam foundation: - Drove a drainage gallery into the downstream foundation. - Installed a system of drainage and observation holes. Checked that the drainage was efficient. Checked the drop in pore pressures. Pore pressures have remained stable ever since
<u>Svartevann Dam</u> Rockfill dam 129m/1976 Zoned dam Moraine core	Documented satisfactory behaviour during construction and operation - Total settlement was somewhat larger than predicted. - Pore pressures in core measured during early construction to check stability: low pore pressures allowed steeper upstream slope than initially designed. - Small leakage.
<u>Storvatn Dam</u> Rockfill dam 90m/1987 Inclined asphalt core	Documented the deformation behaviour of asphaltic core Used the observations to calibrate the analytical models Provided useful information for future dams of this type

[2] Texts in italics are quotes from Peck's (1969).

There is a potential for combining the Observational Method (OM) with the Bayesian updating approach (Christian and Baecher, 2011; Lacasse and Höeg 2019). The OM is a practical way to deal with uncertainty. Bayes' theorem provides a framework that enables updates of first estimates with new information. Bayes' theorem is the essential means of adjusting one's opinion in the light of new evidence. In fact, it is a tool made for geotechnics, as most of what geotechnical engineers do is Bayesian! Most often, the estimates of soil profiles, soil properties, model uncertainties and predictions are based on both measurements and earlier experience and engineering judgment. Bayesian thinking was, for instance, used by Alan Turing in solving the German *Enigma* code during WWII (the movie *The Imitation Game*).

Two sets of data (or predictions), in this case the mean value and the standard deviation, can be combined by Bayes theorem, assuming both datasets are normally distributed, to yield an updated estimate:

$$\mu_{updated} = (\mu_1/\sigma_1^2 + \mu_2/\sigma_2^2)/(1/\sigma_1^2 + 1/\sigma_2^2) \quad (2)$$

$$\sigma_{updated} = (\sigma_1^2 \cdot \sigma_2^2)/(\sigma_1^2 + \sigma_2^2) \quad (3)$$

where μ_1 and σ_1 are the mean and standard deviation of the first estimate (prior), μ_2 and σ_2 are the mean and standard deviation of the measurements (likelihood/new knowledge), and $\mu_{updated}$ and $\sigma_{updated}$ are the updated (posterior) estimates of the mean and the standard deviation. The result is an updated average weighted by the inverse of the standard deviations.

The Observational Method and the field monitoring during dam operation could be "complemented" with a Bayesian updating formulation in the assessments. In this way, one could associate uncertainties and outcomes with probabilistic estimates (probability of occurrence and consequences) and quantify the scenarios for making decisions. A dynamic updating of the risk picture (means and standard deviations) with the help of continuous real-time measurements and prepared response scenarios would be an easy way to make designs safer and provide support for "risk-informed" decision-making.

Bayesian updating has been applied to continuously update the latest knowledge of the unknown parameters with the knowledge of new observations. For dams, two examples of successful applications of the Bayesian updating approach are: (1) in an uncertainty analysis of overtopping of a flood mitigation dam, Michailidi and Bacchi (2017) improved information on the flood peaks from historical observations by incorporating supplementary knowledge from different sources, including their associated uncertainty and errors; (2) Andreini *et al* (2019) developed probabilistic models to predict the internal erosion rate in embankment dams. They did reliability analysis of earth dams in terms of the critical shear stress and a coefficient of internal erosion. The Bayesian updating approach was used to quantify the uncertainty of the uncertain model parameters on the basis of observations in situ jet erosion tests.

Folayan *et al* (1970) were the first to introduce the application of Bayes' theorem to geotechnical engineering. They used Bayesian updating to predict the settlements of a marshland development analysed the associated economic consequences. The approach was also used to illustrate the optimal number of samples to improve the reliability of the prediction. Figure 7 illustrates the overconfidence that can occur in prior subjective estimates of probability distributions, in this case for the compressibility of San Francisco Bay mud.

10 INSIGHT FROM STATISTICAL AND RELIABILITY ANALYSES

For each of the case studies herein and other case studies in Norway and other countries, the reliability analyses have provided useful, but different, insight and formation. For each of the cases reported herein:

- For Dravladalen Dam, the failure mode screening and reliability considerations led to the identification of a so far unidentified, but significant, failure mode. Rehabilitation was required. The analyses documented the important risk reduction through the rehabilitation.
- For Nyhellervatn Dam, the continuous leakage monitoring provided considerable information that confirmed no indication of internal erosion. Nyhellervatn Dam is perceived as solid, robust and well-behaved. Consideration of the failure probability of the downstream slope suggests that there is no need for rehabilitation, even if the traditional deterministic analysis suggest the need for rehabilitation.
- The analyses of the Nesjen Main Dam also suggest a safe and robust dam. Internal erosion is the critical failure mechanism. Rehabilitation measures may further reduce the risk. The analyses show that an optimal solution is achieved if one plans for controlled overtopping of Secondary Dam 4 in the case of an extreme flood event. Overtopping of Secondary Dam 4 has relatively smaller consequences, and water discharge would reduce considerably the risk of a breach at the Main Dam.

Another rockfill dam in Norway, which had been subjected to internal erosion during its first 20 years (the dam is now 50 years old), looked into the failure probability associated with further internal erosion and possible overtopping due to a massive rock slide in the dam reservoir. The analyses quantified the risk reduction potential of different rehabilitation measures and documented that the most expensive measures are not necessarily the most risk-reducing ones.

An additional benefit of reliability-based approaches, although not illustrated herein, lies in the fact that the process of calculating probability of failure reveals which uncertainties are the most significant for the calculated failure probability, and which are unimportant. This additional information provides an effective guide to what improvements in knowledge will reduce the overall uncertainty and failure probability. Although the number calculated is never highly precise, the range of failure probabilities provide a valuable supplement to other measures of safety. For this reason alone, calculating reliability index and/or failure probability adds value to geotechnical engineering analyses

11 SUMMARY AND CONCLUSIONS

This paper aimed at illustrating that a reliability-based approach is more rigorous and more "complete" than the deterministic approach alone because it accounts for the uncertainty in the analysis parameters, their correlation, and therefore leads to a more robust design, and a more rational assessment of safety of existing structures. The reliability-based approach is not meant to replace the traditional deterministic approach. Instead, it should be used as a complement to deterministic analyses. Examples illustrate the complementary information and value added of probabilistic analyses. Recommendations are made for acceptable risk level and on exploiting the seminal Observational Method during dam construction and drawing benefit from continued monitoring during dam operation.

An analysis that allows for both deterministic and probabilistic modelling provides an improved understanding of the potential range of behaviour under various uncertainties. The design criterion in reliability-based design is best defined in terms of a target annual reliability index or failure probability.

There is an increasing demand to adopt a reliability-based approach for the design of geotechnical structures. The techniques have considerably evolved in past years, and the probabilistic analysis software packages are accessible and easy to use. Probability and risk concepts have now reached maturity for our profession. Most owners and operators understand the concept of risk. They expect engineers to provide probability and risk information and/or quantification to help them make risk-informed decisions.

A reliability-based estimate is not necessary for all dam engineering problems, but such estimates should be used to a greater extent when there are significant uncertainties that could influence the result or lead to decisions other than those based on deterministic results alone. All levels of reliability analyses qualitative or quantitative, are useful to gain insight and help make better engineering and management decisions.

There is today a cultural shift in civil engineering practice and the social perception of the engineer's work. Table 10 compares the earlier focus (first column) with the new directions (second column).

Table 10. Cultural shift in civil engineering practice

Earlier focus	To new direction
Hazard	Consequence
Response	Preparedness and Risk Reduction
Reactive	Proactive
Science-driven	Multi-disciplinary
Response Management	Risk Management
Planning *for* communities	Planning *with* communities
Decision-making	Risk-informed decision-making

Science and engineering help us to predict hazards and their probability of occurrence. Knowing the hazards and the risk helps making risk-informed decisions. The time to implement the tools that enable us to qualify or quantify the risk is now. In the future, the dam engineering profession will need to show that the key decisions were "risk-informed" (ISO2394:2015).

12 ACKNOWLEDGMENTS

The authors thank Energi Norge for their support of the analyses. The analyses would not have been possible without the cooperation of Reidar Birkeland, Statkraft AS, Guttorm Mathismoen, E-CO Energi, Rolv Guddal, Sira Kvina, and the participants in the workshops.

13 REFERENCES

Andreini, M., Gardoni, P., Pagliara, S., Sassu, M. 2019. Probabilistic models for the erosion rate in embankments and reliability analysis of earth dams. *Reliability Engineering and System Safety*. 181: 142–155.

Ang, A, H-S., Tang, W.H. 2007. *Probability concepts in engineering. Emphasis on applications to civil and environmental engineering.* 2nd Ed., Wiley and Sons. 406 pp.

Baecher, G. B., Christian J. T. 2003. *Reliability and Statistics in Geotechnical Engineering.* Chichester, West Sussex, England; Hoboken, NJ, J. Wiley. 618 pp.

Benjamin, J. R., Cornell, C. A. 1970. *Probability, statistics, and decision for civil engineers*. McGraw-Hill, New York. 704 pp.

Christian J.T., Baecher G.B. 2011. Unresolved Problems in Geotechnical Risk and Reliability. State-of-the-Art. ASCE Georisk 2011. Atlanta https://doi.org/10.1061/41183(418)3.

EBL 2003/NGI 2002. Stability and Breaching of dams. Shear Strength of Rockfill and Stability of Dam Slopes. Oslo. Publ.123-20033.

FEMA 2015. Federal Guidelines for Dam Safety Risk Management. Federal Emergency Management Agency. P-1025.

Folayan, J.I., Hoeg, K.. Benjamin, J.R. 1970. Decision theory applied to settlement predictions. ASCE *Journal of Soil Mechanics and Foundation*. **96**(4): 1127–1141.

Geotechnique 1996. *The observational method in geotechnical engineering.* The Inst. of Civil Engineers. Thomas Telford London. 332 pp.

Hartford, D.N.D., Baecher G.B. 2004. *Risk and uncertainty in dam safety*. T. Telford. 391 pp.
Høeg, K. 1996. Performance Evaluation, Safety Assessment and Risk Analysis for Dams. *Hydro-power and Dams*. **6** (3): 8 pp.
Høeg, K., Murarka, R.P. 1974. Probabilistic analysis and design of a retaining wall. *ASCE JGGE*. Proc. **100** (GT3): 349−366.
ICOLD 1993. *Rock Foundations for Dams*. ICOLD Bulletin 88.CIGB. Paris. 241 pp.
ISO 31000:2018. Risk management. Guidelines. ISO 31000 2nd edition 2018-02. 16 pp.
ISO 2394:2015. General principles on reliability for structures. ISO/TC 98/SC 2. 111 pp.
Jamiolkowski, M. 2014. Soil Mechanics and the observational method: challenges at the Zelazny Most copper tailings disposal facility. *Géotechnique*. **64** (8) 590–619.
Jamiolkowski, M, Carrier, WD, Chandler, RJ, Høeg, K, Swierczynski, W, Wolski, W 2010. The geotechnical problems of the second world largest copper tailings pond at Zelazny Most, Poland. Proc. 17th SEAGC Southeast Asian Geotech. Conf. Taipei, Taiwan (eds J.C.C. Li and M.L. Lin). **2**: (12–27).
Johansen, P.M. Vick, S.G., Rikartsen, C. 1997. Risk analyses of three Norwegian rockfill dams. Proc. Intern. Conf. on Hydropower. Hydropower'97. Trondheim. pp. 431−442.
Keaveny, J.M., Nadim, F., Lacasse, S. 1990. Autocorrelation functions for offshore geotechnical data. ICOSSAR 1990. Intern. Conf. Structural Safety and Reliability. Perth. Australia. 263−270.
Lacasse, S., Nadim, F. 1996. Uncertainties in characterizing soil properties. *Uncertainty in the Geologic Environment: From Theory to Practice (Uncertainty '96)*, ASCE GSP **158** (49−75).
Lacasse, S., Nadim, F. 2007. Probabilistic geotechnical analyses for offshore facilities. *Georisk*, **1**(1): 21−42.
Lacasse, S., Nadim, F., Liu, Z.Q., Eidsvig, U.K. 2017. Reliability Analysis of Embankment Dams. Geotechnical Frontiers 2017: Geotechnical Materials, Modelling and Testing. ASCE GSP **280** (910−920).
Lacasse, S, Hoëg, K. 2019. In praise of monitoring and the Observational Method for increased dam safety, ICOLD 2019 Annual Meeting/Symposium. Ottawa. 12 pp.
Langford, J. Norén-Cosgriff, K, Lacasse, S., Soergjerd, T., Brekke, A., Koelsch, G. 2019. Risk assessment of vibrations caused by groundworks. 19th ECSMGE. Reykjavik, Iceland.
Leps. T.M. 1970. Review of shearing strength of rockfill. *ASCE Journal of Soil mechanics and Foundation Engineering*. 96(4): 1159−1170.
Liu, Z.Q., Nadim, F., Garcia-Aristizabal, A., Mignan, A., Fleming, K., Luna, B. Q. 2015. A three-level framework for multi-risk assessment, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards.* **9**(2): 59−74.
Michailidi, E.M., Bacchi, B. (2017). Dealing with uncertainty in the probability of overtopping of a flood mitigation dam. *Hydrology and Earth System Sciences*. **21**(5): 2497–2507.
Morgenstern, N. R. 1995. Managing risk in geotechnical engineering. 10th Panamerican Conf. Soil Mechanics and Foundation Eng, **4**:102–126. Guadalajara, Mexico.
Nadim, F., Lacasse, S. 1992. Probabilistic Bearing Capacity Analysis of Jack-Up Structures. *Can. Geotech. J.*, **29**, 580–588.
Peck, R.B. 1969. Advantages and Limitations of the Observational Method in Applied Soil Mechanics. *Géotechnique*. **19**(1): 171–187.
Schneider, J. 1997. Introduction to Safety and Reliability of Structures, Intern. Assoc. for Bridge and Structural Engineering. Structural Engineering documents **5**. 138 pp.
Tang, W.H. 1973. Modelling, Analysis and Updating of Uncertainties. *Proc., ASCE National Meeting on Structural Engineering*, San Francisco.
Tang, W.H. 1984. Principles of probabilistic

characterization of soil properties. *Probabilistic Characterization of Soil Properties: Bridge Between Theory and Practice,* ASCE, 4−89.

Terzaghi, K. 1929. Effect of minor geologic details on the safety of dams. American Inst. Mining and Metallurgical Eng. **215**: 31–44.

Terzaghi, K. 1961. Engineering geology on the job and in the classroom: Past and future of applied soil mechanics. Harvard soil mechanics series. **62**.

US EPA. 2010. Defining Life Cycle Assessment (LCA) US Environmental Protection Agency. 17 October 2010. Web.

USACE 2011. US Army Corps of Engineers. Engineering and Design. Safety of Dams – Policy and Procedures. Engineer Regulation ER 1110-2-1156. Washington DC. USACE. http://publications.usace.army.mil/publications/eng-regs/ER_1110-2 1156/ER_1110-2-1156.pdf.

USACE 2014. Engineering and Design. Safety of dams – policy and procedures, Department of the Army, ER 1110-2-1156, US Army Corps of Engineers, Washington, DC 20314-1000. 31 March 2014.

US Bureau of Reclamation 2011. Dam Safety Public Protection Guidelines. Dam Safety Office, Denver, Colorado. August. http://www.usbr.gov/ssle/damsafety/documents/PPG201108.pdf

Uzielli, M., Lacasse, S., Nadim, F., Phoon, K.K. 2007. Soil variability analysis for geotechnical practice. In Tan *et* al (eds.), Proc. 2nd Intern. Workshop Characterization and Engineering Properties of Natural Soils. **3**:1653−1752. Singapore, Taylor and Francis.

Vick, S. 1994. Geotechnical risk and reliability-From theory to practice in dam safety. Proc., Earth, Engineers and Education. Symp. in Honor of R.V. Whitman, MIT, Cambridge, MA. 45–58.

Vick, S. 2002. *Degrees of Belief. Subjective Probability and Engineering Judgment*. ASCE Press. 405 pp.

Whitman, R.V. 1984. Evaluating calculated risk in geotechnical engineering. *ASCE JGE* **110**(2): 143−188.

Yang, B.B. Yin, K.L., Lacasse, S., Liu, Z. 2019a. Time series analysis and long short-term memory neural network to predict landslide displacement. *Landslides*. Journal of the International Consortium on Landslides. doi:10.1007/s10346-018-01127-x.

Yang, B.B., Yin, K.L., Liu, Z., Lacasse, S. 2019b. Machine learning to predict landslide displacement in dam reservoir. Proc ICOLD 2019. Eds J.P. Tournier *et al*. International Commission on Large Dams. Symposium, Sustainable and Safe Dams Around the World. Paper #485. CRC Press. 8 pp.