Identifying risky scenarios in nuclear waste management using Bayesian networks

Peter Nybergh

School of Science

Bachelor's thesis
Espoo 14.1.2021

Supervisor

Prof. Ahti Salo

Advisor

MSc Edoardo Tosoni

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Peter Nybergh

Title Identifying risky scenarios in nuclear waste management using Bayesian networks

Degree programme Kandidatprogrammet i Teknikvetenskap

Major Matematik och systemvetenskap **Code of major** SCI

Teacher in charge Prof. Ahti Salo

Advisor MSc Edoardo Tosoni

Date 14.1.2021 **Number of pages** 19 **Language** English

Abstract

Conventional risk importance measures are ill suited for more complex systems, where the studied components do not have clearly defined failure states. This work studies two of the recently defined risk importance measures for scenarios, namely the risk achievement worth (RAW) and the risk reduction worth (RRW). The approach is based on a definition of scenario that allows to attach measures of risk importance to individual as well as combined components' states with synergetic or antagonist interactions from the viewpoint of safety.

These RIMs are studied in the context of a case study of a near-surface repository based in Dessel, Belgium. The results show that these RIMs consistently identify risky and safe scenarios. They also confirm relations that exist between the different RIMs. This work serves as a proof of concept for these RIMs, and discusses the need for further development, in the form of extending more conventional RIMs to scenarios.

Keywords scenario analysis, Bayesian networks, risk importance measures, nuclear waste

Aalto University
School of Science

Aalto-universitetet, PB 11000, 00076 AALTO
www.aalto.fi
Sammandrag av kandidatarbetet

Författare Peter Nybergh

Titel Identifiering av riskfyllda scenarier i kärnavfalls hantering med hjälp av Bayesiskt nätverk

Utbildningsprogram Kandidatprogrammet i Teknikvetenskap

Huvudämne Matematik och systemvetenskap **Huvudämnets kod** SCI

Ansvarslärare Prof. Ahti Salo

Datum 14.1.2021 **Sidantal** 19 **Språk** Engelska

Sammandrag

Nuvarande former av riskanalys lämpar sig dåligt för undersökningen av mera komplexa system, där de undersökta variablerna inte har exakta feltillstånd, och där växelverkan mellan variablerna kan påverka betydligt hur en variabel uttrycker sig i systemet. Därför har man utvecklat en ny definition för tidigare riskmått, så att de kan användas tillsammans med scenarier, för så kallad scenarieanalys. Ett exempel på ett sådant system är ytnära förvaringsplatser för kärnavfall, vilka är i fokus för fallstudien i detta arbete. Syftet med detta arbete är att identifiera de riskfyllda scenarier relaterade till ytnära förvaringsplatser av kärnavfall, med hjälp av dessa nyutvecklade riskmått. Detta innebär att identifiera de kombinationer av variabeltillstånd, som gör att det använda riskmåtten antar sina största värden. Bayesiska nätverk ett vanligt sätt i sådana här sammanhang att simulera relationerna mellan de olika variablerna i systemet, och används även i detta arbete.

I detta arbete analyseras användningen av två riskmått, nämligen riskuppnåelsevärde och riskreduceringsvärde. Det förstnämnda riskmåttet beskriver den proportionella förändringen i förväntade risken som uppkommer då man antar att ett visst scenario sker. Det andra riskmåttet å sin sida, beskriver den proportionella förändringen i förväntade risken som uppkommer då man antar att scenario inte sker.

Dessa riskmått används konkret i samband med en fallstudie, där data kommer från ett ytnära förvaringställe i Belgien. Riskmåtten beräknas med hjälp av en kod som utnyttjar bayesiska nätverk samt datan från fallstudien. Resultaten visar bland annat att riskuppnelsevärdet och riskireduceringsvärdet håller med varandra om vilka scenarier som är farliga och vilka scenarier som är säkra, dock så rankar de scenarierna i olika ordningar. De håller alltså inte alltid med om vilket scenarie som är mest riskfyllt för systemet. Med resultaten kan man även motivera varför användningen av olika riskmått är nödvändig, då de berättar på olika sätt om systemets risk, samt hur man bäst kan reducera risken. Resultaten visar också varför det kan vara värt att räkna risken för scenarier och inte för enskilda komponenter, för då flera komponenter granskas på samma gång som en del av ett scenarie får man en bättre bild på växelverkan mellan systemets olika komponenter.

Nyckelord scenarie analys, bayesiska nätverk, riskmått, kärnavfall

Contents

1 Introduction

Nuclear power may lead to lower emissions of greenhouse gases compared to fossil fuels, but it produces hazardous waste of considerable longevity. Consequently, nuclear waste repositories should be designed and built to last over several millennia [Säteilyturvakeskus, 2015]. Such an extended time horizon, along with the complexity of the interacting physical and chemical factors, entails large uncertainty about the evolution of the repository and its surrounding environment [Tosoni et al., 2018].

This uncertainty makes it challenging to assess the safety of nuclear waste repositories. Thus, multiple scenarios can be generated and analysed to characterize the risk due to the radiological impact of the repository. Particularly, Bayesian networks have recently been considered for estimating risk as the probability that the radiological impact violates a reference safety threshold [Salo et al., 2019a].

Nevertheless, although relevant to the safety assessment, the aggregated risk estimate does not help identify which scenarios contribute to risk more than others. From this perspective, useful insights can be gained through risk importance measures [Beeson and Andrews, 2003; Noroozian et al., 2018; Zio, 2011].

However, conventional risk importance measures cannot be straightforwardly applied to Bayesian networks for nuclear waste repositories because i) they require the ex ante classification of the individual system components' states as "functioning" or "failed" from the viewpoint of impairing the overall system's safety, and ii) they do not highlight the effects on risk of the causal dependences between these components.

This work examines how risk importance measures can be extended to scenarios to overcome these limitations. Specifically, we identify and rank the most important scenarios in the case study of an illustrative nuclear waste repository. The rankings resulting from using different risk importance measures are compared and discussed.

The work is structured as follows: Section 2 outlines the background of the work and points out the methodological limitations in the identification of the most important scenarios. Section 3 extends the definitions of some conventional risk importance measures to scenarios, and presents optimization problems for calculating these measures. In Section 4, results are obtained and discussed with reference to the above-mentioned case study. Finally, Section 5 summarizes the work by drawing conclusions on the benefits and limitations of risk importance measures for scenarios.

2 Background

In general, risk importance measures help identify those components of the system whose performance is particularly important for the safety of the system. This can be, for instance, because the failure of such components would increase the probability of system failure to levels which are unacceptable, or because ensuring the adequate performance of these components would improve the safety of the system significantly.

Risk importance measures have been primarily formulated in the context of fault trees [Zio, 2011]. These are logical representations of systems consisting of components with binary states and whose behaviour is governed by boolean logic. The overall performance of the system, too, is assessed in binary terms of success vs. failure. Different risk measures exist which focus on different quantities related to risk (e.g., probability of an event occurring, probability of an occurred event leading to an unacceptable consequence).

However, in the case of *non-coherent* systems, the safety performance of the system can be impaired also by functioning states of some of its components. [Beeson and Andrews, 2003]. This suggests that it can be more meaningful to associate risk importance measures with the *states* of components rather than the components themselves (given that it cannot be taken for granted that the risk would be caused by the failure of the component).

In more general terms, while the notion of component failure is applicable to technical components such as pumps, valves and heat exchangers, this is not the case for more complex techno-environmental systems. For example, the chloride concentration in the deep geological disposal of nuclear waste can be measured by introducing discrete states which correspond to different concentration levels. Then, low concentration levels can cause erosion of the bentonite buffer around the the waste canisters, while high concentration levels may contribute to the corrosion of the copper overpack of the canisters. In such a situation a supposed failure state of chloride concentration would not have a clear interpretation. Rather, conclusions about which concentration levels pose more risks can be reached only through systemic quantitative analysis.

3 Model

3.1 Scenarios in Bayesian networks

Bayesian networks (BNs)[Pearl and Russell, 2003] are probabilistic causal models in which a system is represented as a set V of nodes and a set $A = \{(j,i)|i,j \in V, i \neq j\}$ of directed arcs so that $(j,i) \in A$ indicates that node i depends on node j. We assume the BN to be *connected*, i.e., for any pair $k,l \in V$ there is a sequence of nodes $(i_1,...,i_j,...,i_n), 1 \leq j \leq n$, such that $i_1 = k, i_n = l, (i_j, i_{j+1}) \in A \vee (i_{j+1}, i_j) \in A$, and *acyclic*, i.e., there is no sequence $(i_1,...,i_j,...,i_n), 1 \leq j \leq n, (i_j, i_{j+1}) \in A$, such that $i_1 = i_n$.

Any node $i \in V$ with arcs pointing towards it belongs to the set of dependent nodes $V^D = \{i|i \in V, V_-^i \neq \varnothing\}$. Here, $V_-^i = \{j|(j,i) \in A\}$ is the set of *parents* of i, which is therefore called their *child*. Complementarily, the nodes without parents form the set $V^I = \{i|i \in V, V_-^i = \varnothing\}$ of independent nodes.

Each node is associated with a random variable $X_i, i \in V$ with discrete states $s_i \in S_i$ such as low, medium and high. An independent chance node $i \in V^I$ assumes the state s_i with probability p_{s^i}. For a dependent node $i \in V^D$, the state probability $p_{s^i|s_-^i}$ is conditioned on the combination $s_-^i \in S_-^i = \times_{j \in V_-^i} S_j$ of states of its parents (where $\times$ denotes the Cartesian product).

Let also $s = \{s_i\}, \forall i \in V$, denote a specific combination of states of all chance nodes, and refer to it as a full *path* in the BN. The set of all such paths is, then, $S = \times_{i \in V} S_i$. According to the *global semantics* of BNs [Pearl and Russell, 2003], the probability of a path is

$$p(s) = \prod_{i \in V^I} p_{s_i} \cdot \prod_{j \in V^D} p_{s_j|s_{V_-^j}}, \tag{1}$$

where s_i, s_j and $s_{V_-^j}$ are the states of i, j and V_-^j as specified by s, respectively. These probabilities define the full joint probability distribution function over the realizations of the random variables $X_i, i \in V$, associated with the chance nodes of the BN.

Additionally, let U be the set of values nodes, which are the sink of the network in that they have no children, i.e., $(u,i) \notin A, \forall u \in U, \forall i \in V$. The outcome at a value node is a function $X^u : S_-^u \mapsto \mathbb{C}^u, \forall u \in U$, of its parents' states. Because a value node serves to assess safety, there also exists a real-valued function $\mathcal{U} : \mathbb{C}^u \mapsto \mathbb{R}, \forall u \in U$, of its outcomes. We assume $\mathcal{U}$ to be nonnegative and, as usual in decision theory, unique up to positive affine transformations. More specifically, $\mathcal{U}$ is a *disutility* function such that large values represent less desirable consequences on, e.g., the public or the environment.

Then risk corresponds to the expected disutility. If there is but one value node μ, i.e., $|U| = 1$, the risk implied by the system is

$$\mathbb{E}[\mathcal{U}(X^u)] = \sum_{s \in S} p(s) \cdot \mathcal{U}[X^u(s_{V_-^u})]. \tag{2}$$

By comparison with a predefined risk limit, this estimate helps assess whether the system is safe or not.

Nonetheless, it is also useful to identify the scenarios that are most important in determining the overall risk. To this end, we interpret a scenario as a set of paths. Formally, we define the set $\mathcal{S}$ of all scenarios $\mathbb{S}$ as the powerset $\mathcal{P}(S)$ of S, excluding the empty set (where by $\mathbb{S} \neq \emptyset$) and S itself (where by $\mathbb{S} \neq S$) to prevent vacuous and noninformative scenarios, respectively.

3.2 Risk importance measures

This work investigates two of the recently proposed risk importance measures for scenarios [Salo et al., 2019b]. These measures are extended from system components to scenarios defined as sets of paths, which are combinations of component states. This extension is compatible with conventional risk importance measures in that it is still possible to examine the importance of individual components' failures. Quite importantly, though, this approach makes it possible to measure risks associated with states of components for which failure is not defined, and which are related by uncertain casual dependencies.

As a risk measure, the *risk achievement worth* (RAW)

$$\text{RAW}(\mathbb{S}) = \frac{\mathbb{E}[\mathcal{U}(X_u)|\mathbb{S}]}{\mathbb{E}[\mathcal{U}(X_u)]} \tag{3}$$

is the relative change in the expected disutility from the baseline in case scenario $\mathbb{S}$ occurs. This ratio is well defined unless all paths are associated with consequences which have null disutilities (in which case the baseline level is zero).

If $\text{RAW}(\mathbb{S}) \geq 1$, the occurrence of the scenario increases the risk level, which implies that scenario $\mathbb{S}$ is risky.

As a risk measure, *risk reduction worth* (RRW)

$$\mathrm{RRW}(\mathbb{S}) = \frac{\mathbb{E}[\mathcal{U}(X_u)]}{\mathbb{E}[\mathcal{U}(X_u)|\overline{\mathbb{S}}]} \tag{4}$$

is the relative change in the expected disutility in case the scenario $\mathbb{S}$ does *not* occur (which implies the occurrence of its complement, $\overline{\mathbb{S}}$). Therefore, if $\mathrm{RRW}(\mathbb{S}) \geq 1$, the non-occurrence of scenario $\mathbb{S}$ lowers the risk, which implies that scenario $\mathbb{S}$ is risky. Specifically, the equivalence

$$\mathrm{RAW}(\mathbb{S}) \geq 1 \iff \mathrm{RRW}(\mathbb{S}) \geq 1, \tag{5}$$

indicates that risky scenarios are identified consistently by the two measures.

Furthermore, the relations

$$\mathrm{RRW}(\mathbb{S}_j) \geq \mathrm{RRW}(\mathbb{S}_i) \geq 1, p(\mathbb{S}_j) \leq p(\mathbb{S}_i) \implies \mathrm{RAW}(\mathbb{S}_j) \geq \mathrm{RAW}(\mathbb{S}_i), \text{ and} \tag{6}$$

$$\mathrm{RAW}(\mathbb{S}_j) \geq \mathrm{RAW}(\mathbb{S}_i) \geq 1, p(\mathbb{S}_j) \geq p(\mathbb{S}_i) \implies \mathrm{RRW}(\mathbb{S}_j) \geq \mathrm{RRW}(\mathbb{S}_i) \tag{7}$$

help understand how risky scenarios are ranked by the RAW and the RRW. Particularly, the RRW tends to be low for scenarios with low probabilities, because assuming that they will not occur does not constitute a significant change from the baseline situation. Thus, a scenario with a larger RRW, but with a lower probability than another, must be a riskier scenario that has a larger RAW, too (6). Analogously, the RAW is small for scenarios with high probability, because their occurrence would not imply large differences from the expectations (unless their consequences have very large disutilities). Hence, if a scenario has a larger RAW and a higher probability than another, then it must also have a larger RRW (7).

3.3 Calculating RIMs through optimization

A challenge in analysing risk importance with scenarios is that the number of scenarios grows rapidly with the number of components and their states. For example, if there are five components with three states for each, there are $3^5 = 243$ paths from which one can generate $2^{243} - 2 \approx 1,4 \times 10^{73}$ different scenarios (excluding the two cases mentioned in Section 3.1). The explicit enumeration of all scenarios may consequently be impossible, making it necessary to develop efficient ways to identify which scenarios have the highest values of the risk importance measure. For instance the most important scenarios could be identified via optimization.

Specifically, binary variables are defined as $z(\cdot) : S \mapsto \{0, 1\}$ to either include ($z(s) = 1$) or exclude ($z(s) = 0$) each path from the scenario. All scenarios must satisfy the constraints

$$1 \leq \sum_{s \in S} z(s) \leq |S| - 1 \tag{8}$$

to ensure that the set of paths $\{s \in S | z(s) = 1\}$ is a scenario as per the definition in Section 3.1.

Then, riskiest scenarios can be found by optimizing objective functions corresponding to the different measures. The RAW is the largest for the scenario $\mathbb{S}^*_{RAW}$ for which the conditional risk $\mathbb{E}[\mathcal{U}(X_u)|\mathbb{S}^*_{RAW}]$ attains its maximum, so that the optimization problem is

$$\max_{z(s)} \quad \frac{\sum_{s \in S} z(s) p(s) \mathcal{U}[X_u(s_{V^u_-})]}{\sum_{s \in S} z(s) p(s)}. \tag{9}$$

Instead, as it can be shown [Salo et al., 2019a], the RRW can be maximized by minimizing the RAW of the complement set $\overline{\mathbb{S}}^*_{RRW}$ which contains all the paths that are *not* in the riskiest scenario $\mathbb{S}^*_{RRW}$ being searched for. Therefore, the optimization problem is

$$\max_{z'(s)} \quad \frac{\sum_{s \in S} z'(s) p(s) \mathcal{U}[X_u(s_{V^u_-})]}{\sum_{s \in S} z'(s) p(s)}, \tag{10}$$

where the binary variable $z'(s) = 1 - z(s)$ (which is required as a constraint in the optimization) is utilized to select and discard the paths which are excluded from and included in $\mathbb{S}^*_{RRW}$, respectively. The challenging nonlinearities implied by the ratios in the objective functions can be eliminated by linear fractional transformations through which the optimization problems (9) and (10) become easily solvable mixed integer linear programs [Salo et al., 2019a].

In both optimization problems, the respective riskiest scenarios $\mathbb{S}^*_{RAW}$ and $\mathbb{S}^*_{RRW}$ are identified by the set of paths $s \in S$ for which $z(s) = 1$ (or $z'(s) = 0$, in the case of RRW). The RAW of $\mathbb{S}^*_{RAW}$ is given by the ration between the optimized objective function in (9) and the baseline risk. Vice-verse, the RRW of $\mathbb{S}^*_{RRW}$ is the ratio between the baseline risk and the optimized objective function in (10).

Different kinds of scenarios can be generated by introducing constraints. Here, we focus on *projected* scenarios, which are defined by restricting the set $S'_i \subsetneq S_i$ of states for one or more nodes in $I \subset V$. Binary variables associated with these nodes' states can be defined such that

$$z_i(s_i) = \begin{cases} 1, & s_i \in S_i' \\ 0, & \text{otherwise.} \end{cases}$$

The size of the sets S_i' can be bounded through

$$\underline{n}_i \leq \sum_{s_i \in S_i} z_i(s_i) \leq \overline{n}_i, \quad \forall i \in I, \tag{11}$$

with $\underline{n}_i \geq 1$ and $\overline{n}_i \leq |S_i| - 1$.

The consistency between the binary variables for paths and those for the nodes' states can be ensured through further linear constraints. On one hand,

$$z(s) \leq \frac{1}{|I|} \sum_{i \in I} z_i(s_i) \tag{12}$$

guarantees that the scenario does not contain paths for which $z_i(s_i) = 0, i \in I$. On the other hand, the constraint

$$z(s) \geq \sum_{i \in I} z_i(s_i) - |I| + 1 \tag{13}$$

implies that the scenario contains all paths for which $z_i(s_i) = 1, \forall i \in I$.

Finally, either optimization problem can be solved repeatedly to find the second, third, etc., riskiest scenarios beside the top one. At each iteration, the constraint

$$\sum_{s \notin \mathbb{S}^*} z(s) + |\mathbb{S}^*| - \sum_{s \in \mathbb{S}^*} z(s) \geq 1 \tag{14}$$

can be added to exclude the scenario $\mathbb{S}^*$ found as the riskiest at the previous iteration. These constraints can be added until sufficiently many of the most important scenarios have been found.

4 Results and discussion

We illustrate the risk importance measures for scenarios through the near-surface repository studied by Salo et al. [Salo et al., 2019b] (Figure 1). In particular, the data refers to the repository planned for the site of Dessel (Belgium). The repository is modeled through a Bayesian network, whose nodes can assume the states in Table

1. These states derive from the discretization of continuous ranges of values for the variables that correspond to the nodes [Uusitalo, 2007]. Their discretized probabilities are found in Table 2.

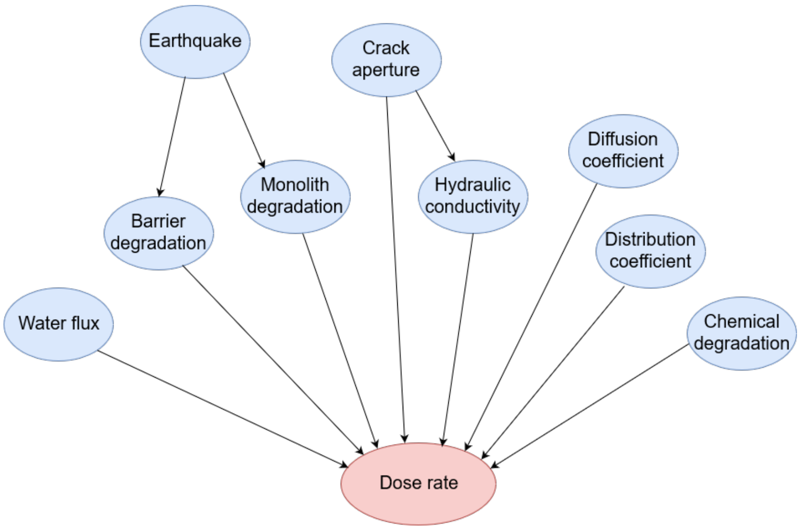


Figure 1: Bayesian network of a near-surface nuclear waste repository.

The states for the FEPs are for example *low* and *high*, or *fast* and *slow*. For *water flux* for example, the states *low* and *high* indicate the speed at which water flows into the system. For *barrier degradation*, on the other hand, the states *fast* and *slow* define the speed at which the protective barrier is degrading.

Table 1: The nodes of the network of Figure 1 and respective states.

FEP	States
Earthquake	BDBE, Earthquake
Water flux	Low, High
Crack aperture	Micro, Macro
Diffusion coefficient	Low, High
Distribution coefficient	Low, High
Chemical degradation	Fast, Slow
Barrier degradation	Fast, Slow
Monolith degradation	Very fast, Fast, Slow
Hydraulic conductivity	Low, Medium, High
Dose rate	Respect, Violation

The FEP *earthquake* is also included, because it can affect the speed of barrier degradation. The arc from *crack aperture* to *hydraulic conductivity* explains the increased effective conductivity of fissured concrete. In this case study, the safety target is the *Dose rate* to the public, normalized by the confidential safety threshold.

Table 2: The probability distributions of FEPs in Figure 3. The numerous conditional probabilities for the *dose rate* are not reported for the sake of brevity.

FEP	Parents	Probability distributions		
		0	1	2
Earthquake	No	0.9954	0.0046	
Water flux	No	0.8641	0.1359	
Crack aperture	No	0.8074	0.1926	
Diffusion coefficient	No	0.5000	0.5000	
Distribution coefficient	No	0.5000	0.5000	
Chemical degradation	No	0.5000	0.5000	
Barrier degradation	Earthquake			
	BDBE	0.0580	0.9420	
	Earthquake	0.3600	0.6400	
Monolith degradation	Earthquake			
	BDBE	0.2950	0.2920	0.4130
	Earthquake	0.2950	0.4250	0.2800
Hydraulic conductivity	Crack aperture			
	Low	0.6670	0.1888	0.1442
	High	0.1888	0.6670	0.1442

We first examine the scenarios made of different states of *hydraulic conductivity*, which are listed in Table 3, along with their RAW, RRW and probabilitiy.

Table 3: The RAW and RRW values for the scenarios of *hydraulic conductivity*.

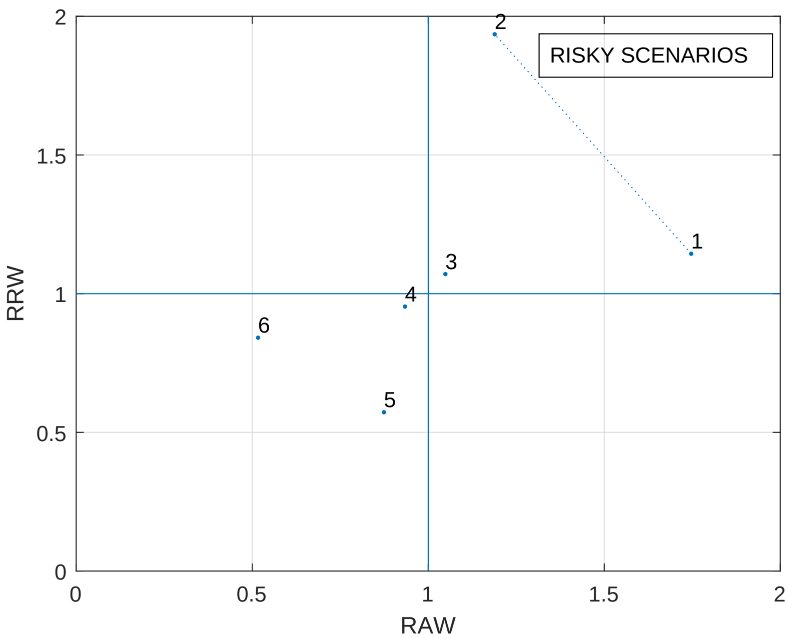
Scenario	Hydraulic Conductivity			RAW	RRW	Probability
	L	M	H			
1			■	1.747	1.144	0.144
2	■		■	1.189	1.935	0.719
3	■			1.049	1.071	0.575
4		■	■	0.934	0.953	0.425
5	■	■		0.874	0.572	0.856
6		■		0.518	0.841	0.281

The scenario where the *hydraulic conductivity* is *high* has the largest RAW, meaning that it should be prevented to avoid an almost two-fold increase in risk. The largest RRW is instead achieved by the scenario in which the *hydraulic conductivity* is either *low* or *high*. This is because the probability for this scenario is significantly larger than for the first scenario. Hence, this scenario should be prevented to achieve the largest risk reduction.

In Table 3, all the scenarios that are risky (i.e., whose RAW and RRW are larger than 1), have the *hydraulic conductivity* either *low* or *high*. Conversely, all scenarios that are safe, are characterized by a *medium hydraulic conductivity*, in some combination with the *low* or *high*. This indicates that risk is not a monotonic function of the states of *hydraulic conductivity*.

Figure 2 confirms, that risky scenarios are consistently identified by RAW and RRW (Relation 5). Furthermore, the figure shows that scenarios 1 and 2 of Table 3 constitute a *pareto front* (dotted line) in that scenario 1 is the most important scenario if the goal is to prevent a significant risk increase during the repository lifetime, while scenario 2 is the most important scenario if the goal is to reduce risk.

Figure 2: RAW and RRW of the scenarios (dots) of *hydraulic conductivity* (numbering as per Table 3). The dotted line represents the *pareto front* of the riskiest scenarios.



Thus, no state can be uniquely identified as the failure state of hydraulic conductivity. These results confirm that nuclear waste repositories are an example of systems for which it is usually not possible to distinguish functioning and failed states of its relevant factors. Rather, each factor's state may be more or less risky depending on the state of the other factors it interacts with.

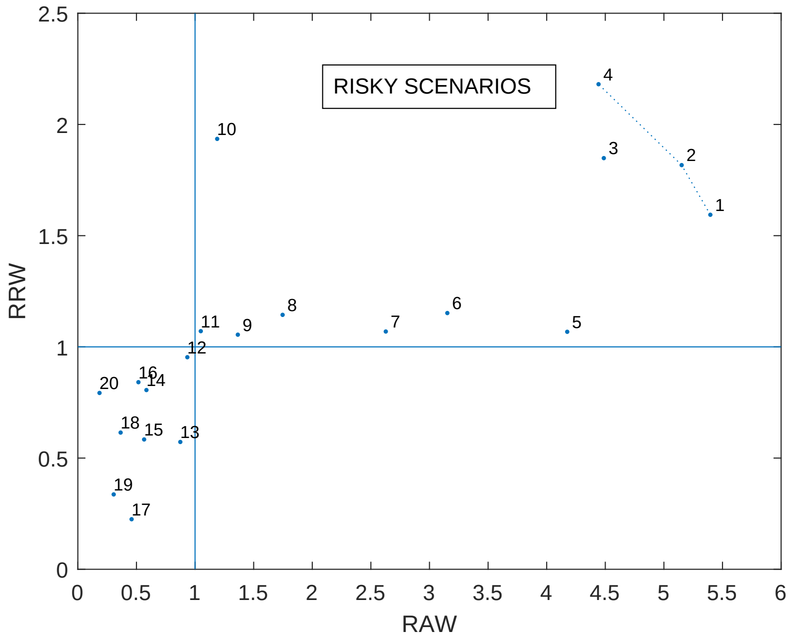
Table 4: The RAW and RRW values for the scenarios of *water flux* and *hydraulic conductivity*. Black squares indicate the possible states for each scenario.

Scenario	Water Flux		Hydraulic Conductivity			RAW	RRW	Probability
	L	H	L	M	H			
1		■	■			5.396	1.594	0.078
2		■	■		■	5.151	1.817	0.098
3		■	■	■		4.487	1.849	0.117
4		■	■	■	■	4.442	2.181	0.136
5		■			■	4.175	1.068	0.019
6		■		■	■	3.152	1.152	0.058
7		■		■		2.627	1.069	0.038
8	■	■			■	1.747	1.144	0.144
9	■				■	1.365	1.055	0.125
10	■	■	■		■	1.189	1.935	0.719
11	■	■	■			1.049	1.071	0.575
12	■	■		■	■	0.934	0.953	0.425
13	■	■	■	■		0.874	0.572	0.856
14	■			■	■	0.585	0.806	0.367
15	■		■		■	0.565	0.584	0.621
16	■	■		■		0.517	0.841	0.281
17	■		■	■	■	0.459	0.225	0.864
18	■		■			0.365	0.615	0.497
19	■		■	■		0.306	0.337	0.739
20	■			■		0.185	0.793	0.243

To investigate these interactions further, let us consider the different scenarios of hydraulic conductivity together with water flux, which are ranked by RAW in Table 4, along with RRW and their respective probabilities. Different from Table 3, *low* is the riskiest state of *hydraulic conductivity* (based on RAW), when matched with *high* water flux. A potential explanation for this might be that, if a large amount of water is seeping through the repository and it encounters the resistance of a low hydraulic conductivity, there might be some accumulation of water and a consequent flooding of the repository. When the state of water flux is unspecified, high hydraulic conductivity gives the largest RAW, but it only ranks eight in Table 4.

The scenarios in Table 4 are also represented in Figure 3.

Figure 3: RAW and RRW of the scenarios (dots) of *water flux* and *hydraulic conductivity* (numbering as per Table 4) .



As can be seen from Figure 3, the determination of the riskiest scenario of all depends on the preference for RAW (scenario 1), RRW (scenario 4) or something between the two (scenario 2). We can also conclude that the scenarios are all correctly identified by both the RAW and the RRW, respectively.

Confirming the relation (Relation 7) scenario 4 has a larger RRW and a lower probability than, for instance, scenario 8, whereby it also has a larger RAW. Similarly, scenario 4 has both larger RAW and probability than, for instance, scenario 5, implying that it also has a larger RRW (Relation 7).

5 Summary

This work has been a *proof of concept* for this new approach presented by Salo et al. [Salo et al., 2019b]. The results illustrate how the new risk importance measures for scenarios enable the analysis of complex, non-coherent and non-binary state systems.

The benefits of this approach it that we can analyse multiple components at a time as well as the effect on risk of single component's states.

The results confirm the relations between the risk importance measures suggested in [Salo et al., 2019b]. Specifically, it has been verified that the RAW and RRW identify the same set of risky scenarios, but may rank them by different risk prioritizations. This means that the riskiest scenario depends on which measure is chosen for the analysis. This is why several different risk importance measures are necessary, as they all give a different angle for the risk analysis. One path for further development could be to extend other risk importance measures to scenarios, and consequently analysing the produced results and their reliability in this new context.

This work focused on one possible application for the RIMs extended to scenarios, namely near-surface nuclear waste repositories, However, other similar systems could benefit from them as well. Another path for further further development could be to have other data sets and other systems as case studies, and show how well they perform in comparison to more conventional methods.

References

S. Beeson and J. D. Andrews. Importance measures for noncoherent-system analysis. *IEEE Transactions on Reliability*, 52(3):301–310, 2003. doi: 10.1109/TR.2003.816397. URL http://dx.doi.org/10.1109/TR.2003.816397.

A. Noroozian, R. B. Kazemzadeh, S. T. A. Niaki, and E. Zio. System risk importance analysis using bayesian networks. *International Journal of Reliability, Quality and Safety Engineering*, 25(1), 2018. doi: 10.1142/S0218539318500043. URL https://dx.doi.org/10.1142/S0218539318500043.

J. Pearl and S. Russell. Bayesian networks. In M. A. Airbib, editor, *The Handbook of Brain Theory and Neural Networks*, pages 157–160. The MIT press, 2003.

A. Salo, J. Govaerts, E. Zio, and E. Tosoni. Comprehensiveness of scenarios in the safety assessment of nuclear waste repositories. *Reliability Engineering and System Safety*, pages 188,561–573, 2019a. doi: 10.1016/j.ress.2019.04.012. URL https://doi.org/10.1016/j.ress.2019.04.012.

A. Salo, E. Tosoni, and E. Zio. Measures of risk importance for probabilistic scenario analysis. Manuscript 2019, 2019b.

Säteilyturvakeskus. Disposal of nuclear waste. STUK-YVL D.5 (version 13.2.2018), Helsinki, Finland, 2015.

E. Tosoni, A. Salo, and E. Zio. Scenario analysis for the safety assesment of nuclear waste repositories: A critical review. *Risk Analysis*, 38(4):755–776, 2018. doi: 10.1111/risa.12889. URL https://dx.doi.org/10.1111/risa.12889.

L. Uusitalo. Advantages and challenges of bayesian networks in environmental modelling. *Ecological Modelling*, 203(3-4):312–318, 2007. doi: 10.1016/j.ecolmodel.2006.11.033. URL http://dx.doi.org/10.1016/j.ecolmodel.2006.11.033.

E. Zio. Risk importance measures. In H. (Eds.), editor, *Safety and Risk Modeling and its Applications*, pages 151–196. Springer-Verlaag, 2011. doi: 10.1007/978-0-85729-470-8_1. URL http://dx.doi.org/10.1007/978-0-85729-470-8_1.