

An extended method for evaluating assumptions deviations in quantitative risk assessment and its application to external flooding risk assessment of a nuclear power plant

Tasneem Bani-Mustafa ⁽¹⁾, Roger Flage ⁽²⁾, Dominique Vasseur ⁽³⁾, Zhiguo Zeng ⁽¹⁾ and Enrico Zio ⁽¹⁾⁽⁴⁾

⁽¹⁾ *Chair on System Science and the Energetic Challenge, EDF Foundation*

Laboratoire Genie Industriel, CentraleSupélec, Université Paris-Saclay,

3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France

⁽²⁾ *University of Stavanger, Norway*

⁽³⁾ *EDF R&D, PERICLES (Performance et prévention des Risques Industriels du parc par la simulation et les Etudes) EDF Lab Paris Saclay - 7 Bd Gaspard Monge, 91120 Palaiseau, France*

⁽⁴⁾ *Energy Department, Politecnico di Milano, Via Giuseppe La Masa 34, Milan, 20156, Italy*

Abstract

In quantitative risk assessment, assumptions are typically made, based on best judgement, conservative, or (sometimes) optimistic judgments. Best judgment and optimistic assumptions may result in failing to meet the quantitative safety objectives, whereas conservative assumptions may increase the margins which the objectives are met with but result in cost-ineffective design or operation. In the present paper, we develop an extended framework for the analysis of the criticality of assumptions in risk assessment by evaluating the risk that deviations from the assumptions lead to a reduction of the safety margins. The framework aims to support risk-informed decision making by identifying important assumptions and integrating the assessment of their criticality into the quantitative risk assessment (QRA). The framework is, finally applied within the quantitative risk assessment of a Nuclear Power Plant (NPP) exposed to external flooding. Compared to previous works on the subject, we consider also conservative assumptions and introduce decision flow diagrams to support the classification of the criticality of the assumptions. The framework provides a more comprehensive and transparent evaluation of the assumptions deviation risk through the decision flow diagrams that facilitate the standardization of the evaluation of the assumption deviation effects on the risk assessment.

Keywords

Quantitative risk assessment; conservative assumption; assumption deviation; strength of knowledge; decision flow diagram; nuclear power plants; external flooding.

1. Introduction

Quantitative or probabilistic risk assessments (QRAs/PRA) have found extensive usage in several industries, starting nuclear power generation (e.g., NRC (1975); Vesely and Apostolakis (1999)), and continuing with, for example, offshore petroleum exploration and production (e.g., Vinnem (2014a, 2014b)) and civil air transport (e.g., Netjasov & Janic (2008)). Broadly speaking, QRA, provides the decision-maker with a quantitative risk description (see Kaplan and Garrick (1981)) for a set of decision options. This risk description serves to support risk-informed decision making, as highlighted by Apostolakis (2004). A number of textbooks present methods and models useful to perform QRAs, e.g., Henley and Kumamoto (1981), Fullwood and Hall (1988), Stewart and Melchers (1997), Bedford and Cooke (2001), Andrews and Moss (2002), Cox Jr (2012), Vose (2008), Zio (2007-2009) and Aven (2011).

In quantitative risk assessments, models are used to represent systems and processes for providing predictions of their safety performance based on given risk metrics (Aven and Zio, 2013). These models are built on a set of assumptions that are translated in mathematical forms for quantitative assessments (Bjerga et al. (2014), NRC (2010), Eiser et al. (2012)). The risk assessments may have a more or less solid foundation, depending on the validity of the hypotheses made (Boone et al. (2010), Klopogge et al. (2011), Berner and Flage (2016)).

Making assumptions is an inevitable part of quantitative risk assessment (QRA) process. An assumption can be defined generally as “a fact or statement (such as a proposition, axiom [...], postulate, or notion) taken for granted” (Merriam-Webster). Other definitions, from the scientific literature and more specific to the risk assessment context, include “conditions/inputs that are fixed in the assessment but which are acknowledged or known to possibly deviate to a greater or lesser extent in reality” (Berner & Flage, 2016 p. 46) and the following, which relies on the definition of defaults (Suter *et al.*, 2007 pp. 134-135):

“Defaults are functional forms or numerical values that are assigned to certain models or parameters in risk assessment, based on guidance and standard practice, in the absence of good data. [...] Assumptions are equivalent to defaults but are derived for a specific assessment rather than being taken from guidance. They may be complex, implying functional forms or sets of parameters. [...] Ad hoc assumptions must be individually justified.”

The latter definition restricts assumptions to having a quantitative format, whereas the former definitions allow also for qualitative assumptions and highlight the potential, or even expected, non-true nature of assumptions. Some examples of types of assumptions in risk assessment are:

1. The number of people exposed to a hazard

2. The reliability of a safety barrier
3. The behavior of people leading up to or following an accidental event.

The first two types of assumptions concern directly quantitative risk model parameters. The last assumption is likely to be more qualitative in nature, e.g. assuming that all people involved in the accidental event follow the emergency preparedness plan. Transforming this qualitatively formulated assumption into a quantitative format is less straightforward.

Risk assessment assumptions are typically of best judgement or conservative. Best judgement assumptions are here understood as reflecting the best knowledge on the matter, e.g. a realistic “best estimate” of a risk model parameter, whereas conservative assumptions come from lack of knowledge on the matter or conscious simplification of its analysis, and define conditions or values that are in some sense ‘unfavorable’, or ‘protective’, with respect to the current knowledge and lack thereof. Optimistic assumptions are also possible, but are typically rare in risk assessment, from the safety perspectives.

For best judgement and optimistic assumptions, deviations of the actual conditions could cause the safety objectives not to be. With regards to this, the concept of assumption deviation risk assessment was coined by Aven (2013) to address this type of “risk” situation by evaluating different intensities of deviations, their associated probabilities of occurrence, the effect of the deviations on the consequences, and an overall strength of knowledge judgement for these three attributes. Assumption deviation risk assessment, thus, goes beyond sensitivity analysis, which tends to be focused on “what if” questions, as discussed by Khorsandi & Aven (2017). In the case of conservative assumptions, on the other hand, deviations might decrease the margins for meeting the objectives.

It is recognized that the risk assessment needs to be enhanced through communicating the assumption deviation risk for a better understanding of the basis upon which the assessment is based. Therefore, a more complete evaluation of the assumption deviation risk is needed. The aim of the current work is to provide a transparent and comprehensive evaluation technique of the assumption deviation risk through a standardized technique, in order to better inform decision-making. In the present paper, we take the recently suggested method for evaluating the risk from assumptions deviations by (Khorsandi & Aven, 2017) and apply it to the external flooding risk assessment of a nuclear power plant (NPP). In doing this, we extend the overall methodology to evaluate also the risk of deviations from conservative assumptions and introduce decision flow diagrams for the quantitative evaluation. We find that the proposed extensions provide a more solid decision making basis than focusing only on best judgement assumptions

and that the decision flow diagrams facilitate a standardization of the evaluation of the risk from assumptions deviations.

Works closely related to the present paper include the already mentioned papers by Aven (2013), introducing the concept of assumption deviation risk, and by Khorsandi & Aven (2017), presenting how to integrate an assumption deviation risk assessment as part of a quantitative risk assessment (QRA). Berner & Flage (2016) also build on the assumption deviation risk concept and develop a framework comprising six classes of uncertain assumptions, which is used to prescribe strategies for treating these assumptions both in the risk assessment (Berner & Flage, 2016) and in the subsequent risk management (Berner and Flage, 2017).

The remainder of the paper is organized as follows. In Sect. 2, we describe the extended method. Then, in Sect. 3, we present the application to the case study. In Sect. 4, we offer a discussion of some conclusions.

2. Extended framework for the evaluation of assumptions deviations

In this section, we extend the original work of Khorsandi and Aven (2017) for a more comprehensive assessment of the criticality (risk) of assumptions deviations. In Sect. 2.1, we present the extended framework and compare it to the original one. In Sect. 2.2, the detailed implementation of the framework is described.

2.1. The assessment framework

In this section, the original work of Khorsandi and Aven (2017) is extended considering multiple contexts of decision-making and multiple types of assumptions. We assume that each assumption As_i affects the numerical values of some parameters in the Probabilistic Risk Assessment (PRA) model. The factor that links the assumptions to the numerical parameters is called “junction” in this paper. The criticality (C) of an assumption is assessed based on the six criteria: (i) the type of assumption; (ii) the context of decision making; (iii) the belief (likelihood) in deviation from reality; (iv) the amount of deviation from reality; (v) the likelihood of the deviation; (vi) the margin of deviation; (vii) the strength of the knowledge supporting the assumption made. Three levels of criticality are defined with their corresponding settings:

1. Very critical ($C = 1$): The assumption is based on weak knowledge and the confidence on the assigned value of the model parameters is low. Besides, the assumption deviation has severe influence on the decision making and might lead to exceedance of the safety limit. Further analysis and justification of the assumption is required.

2. Not very critical ($C = 2$): The assumption is made based on a moderate level of knowledge. The assumption deviation is likely to happen, but the risk metric remains within the safety limits even after considering such assumption deviation. The assumption can be trusted to support decision making if the risks of the deviation from other assumptions are all not critical ($C = 3$). Further analysis and justification of the assumption is needed only when multiple other assumptions are also in this state.
3. Not critical ($C = 3$): The assumption made is based on strong knowledge. An assumption deviation is unlikely to happen or, if it happens, it does not affect the decision making. The assumption can be trusted and decisions can be made based on the current assumption.

To evaluate the criticality of the assumptions deviations, six criteria are considered, as shown in Figure 1:

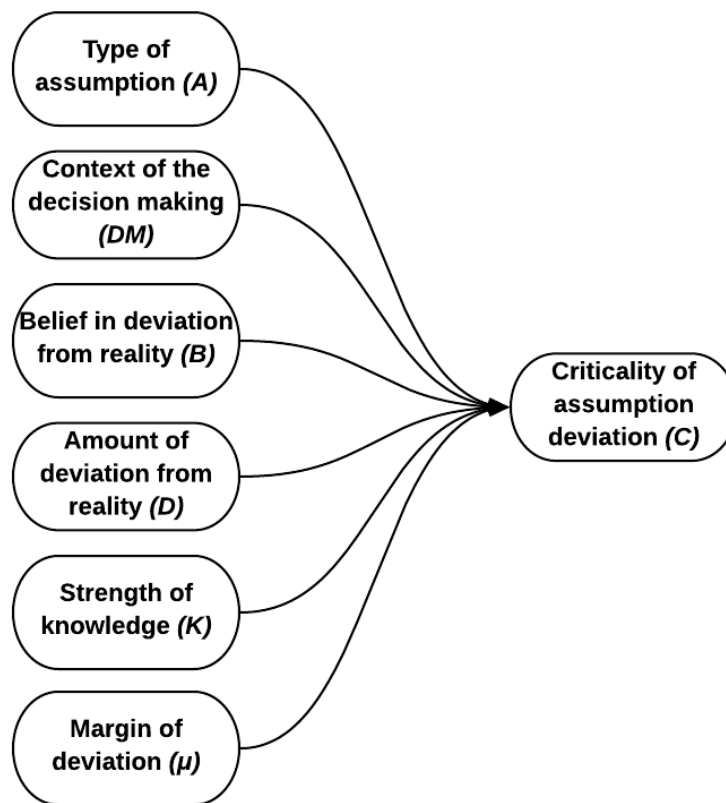


Figure 1 Criteria for evaluating the criticality of assumption deviation.

1. Type of assumption (A): Assumptions made in PRA can be classified into different types. For example, EPRI (2015) distinguishes three types of assumptions: conservative assumptions, best judgment assumptions and approximations. Conservative assumptions are made out of cautiousness and tend to overestimate the risk rather than underestimate it; best judgment assumptions are believed to represent expected scenarios, given the available knowledge; approximations are assumptions that are made for reducing the complexity of the models

(EPRI, 2006). Deviations in different types of assumptions might lead to different influences on the PRA. In our framework, three types of assumptions are considered:

- i. Optimistic assumption (A_1): the assumption is judged by peers to underestimate the risk when compared to reality
- ii. Best judgment (A_2): the assumption is judged by peers as representative of reality (realistic)
- iii. Conservative assumption (A_3): the assumption is judged by peers to overestimate the risk when compared to reality (pessimistic).

Note that unlike EPRI classification, we do not consider approximations as an independent category of assumptions. An approximation can hence be linked to an assumption of any of the three types mentioned above.

2. Context of the decision making (DM): Risk metrics are used to support decision making in different contexts (EPRI, 2015). In this paper, we distinguish between two contexts of decision-making. First, the comparison with safety objectives, whereby the risk metrics are compared to quantitative safety goals and criteria (EPRI, 2015). In this case, the decision maker would accept performing the task (project, task, work, etc.) if the risk metric is lower than the safety objective; otherwise, some safety reduction measures (e.g., safety barriers, safety systems, etc.) need to be implemented in order to reduce the risk. Second, the comparison of alternatives, whereby risk metrics of different alternatives are compared. In this case, the decision maker would choose the alternative that leads to a lower risk, or choose the risk reduction measure that leads to a higher reduction of the risk metric given the cost of the application. The criticality of assumptions deviations varies from one context to another, where, in comparing risk metric to a safety goal, only the deviation toward critical scenarios need to be considered. On the other hand, for comparing alternatives in terms of their risks, all the deviation scenarios need to be considered, since a conservative assumption might lead to a higher risk metric and hence, lead the decision maker to make a wrong decision by choosing another alternative that has a higher risk in reality but appears lower due to the different levels of conservatism in the analysis.
3. Belief in deviation (B) measures the realism of an assumption and is expressed by the likelihood of assumption deviation. The likelihood is assigned by the experts following the criteria defined in Khorsandi and Aven (2017), i.e., what could cause the assumption to deviate in reality; what are the key drivers of those causes; etc.

4. Amount of deviation from reality (D) refers to the amount of deviation between the assumed parameter value and the true value. It is assigned by experts and expressed in percentage.
5. Strength of knowledge (K) refers to the strength of the background knowledge that supports the evaluation of the belief in deviation and the amount of deviation.
6. Margin of deviation (μ) refers to the degree to which an assumption may deviate before the deviation changes the decisions made based on the results of risk assessment, e.g., the violation of the acceptance criteria or the change of the prioritization of different options. This margin is calculated analytically (see Sect. 2.2.8) and expressed in percentage.

The logical combination of the six criteria yields different levels of criticality. Decision flow diagrams are introduced in this paper to capture the logical relationship between the six criteria and the criticality of assumptions deviations (see Sect. 2.2.9). A comparison between the original assessment framework in Khorsandi and Aven (2017) and the extended framework is made in Figure 2. It can be seen that the original work of Khorsandi and Aven (2017) is adjusted and extended to include an additional context of decision making (comparing alternatives) and also a new type of assumption (conservative assumptions). Accordingly, new criteria are added or adjusted to integrate the new decision context and type of assumption in the assessment of the assumption deviation risk. As to the presentation of the assumption deviation risk, the radar plot in Khorsandi and Aven (2017), which presents the contributing factors to the assumption deviation risk individually, is replaced with an overall integrated metric for assumption deviation risk, i.e., the criticality (C). These extensions make it possible for the extended framework to provide a more comprehensive description of the risk from assumptions deviations.

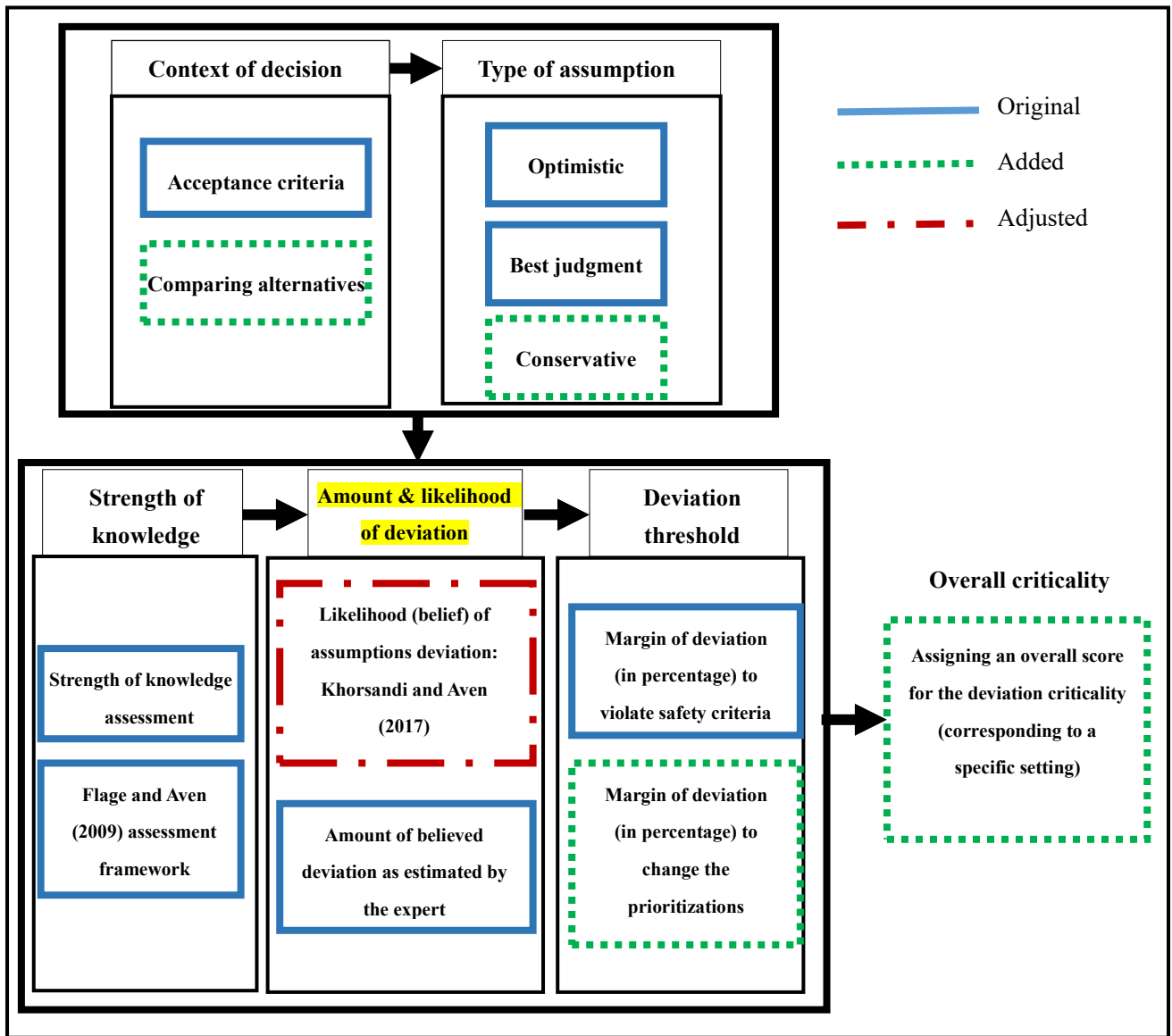


Figure 2 A comparison between the original (Khorsandi & Aven, 2017) and the extended frameworks for assumption deviation risk assessment.

2.2. Implementation of the framework

As shown in Figure 3, nine main steps are needed for applying the developed framework to assess the criticality of assumptions deviations. The nine steps are discussed in details in sub Sect. 2.2.1-2.2.9.

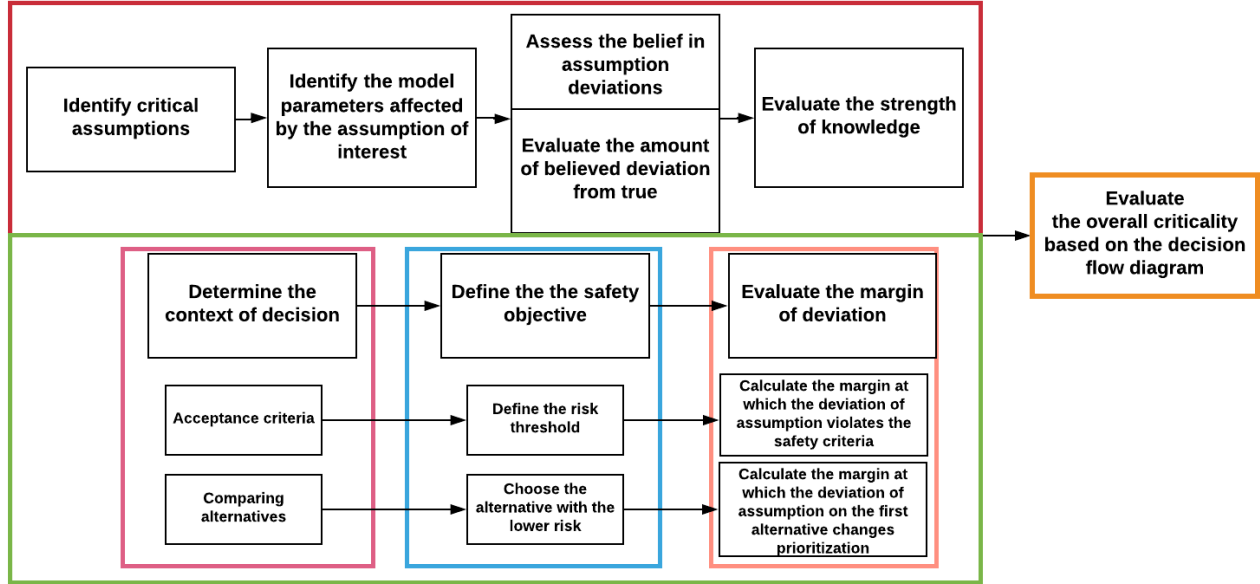


Figure 3 Procedure for applying the developed framework for assumption deviation criticality (risk) assessment.

2.2.1. Identify critical assumptions

In the first step, the assumptions made in the PRA are identified. The assumptions might be made due to lack of understanding and knowledge about a phenomenon or as an attempt to reduce the modeling details and complexity (EPRI, 2006, EPRI, 2015). The type of each assumption (A) is determined by expert judgment, making reference to the definitions in Sect 2.1.

2.2.2. Identify the model parameters affected by the assumption of interest

As mentioned in Sect 2.1, in this paper, we assume that there is a juncture that connects numerically an assumption to one or more parameters in the PRA model. Without losing generality, let us assume that this juncture that the PRA model is represented by:

$$R = f(p_1, p_2, \dots, p_m, \dots, p_n), \quad (1)$$

where R is the risk metric and $p_1, p_2, \dots, p_m, \dots, p_n$ are the model parameters (e.g., failure probabilities), f is the function that depends on the structure of the model. The juncture can be conceptually represented as in Figure 4, where As represents a set of assumptions. In the framework, we only consider the assumptions that can be altered numerically or that can change the numerical values of the model parameters. We do not consider the assumptions

that are related to the model structure or that cannot be measured numerically. The second step, then, involves identifying the model parameters affected by each assumption, as shown in Figure 4.

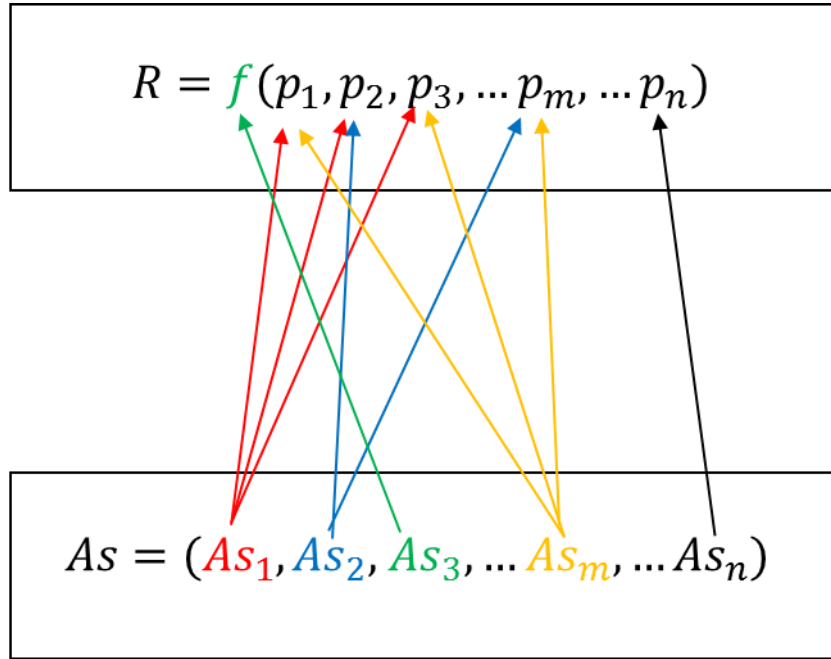


Figure 4 Representation of connections between assumptions and model parameters.

2.2.3. Assess the belief in assumption deviation

The belief in deviation is evaluated as the subjective probability assigned by experts that the assumption deviates from the actual conditions. The assigned value is conditional on the available background knowledge, including experts' individual expertise. It should be noted that the aim of evaluating the belief in deviation is not to assign a precise value for the probability of deviation. Rather, it aims at expressing the experts' beliefs, based on the available knowledge, on how likely the assumption might be deviating from reality (Khorsandi and Aven 2017). Such a step can be regarded as a tool for making good use of experts' individual expertise by reflecting their implicit knowledge that cannot be directly stated or documented.

To determine the value of B , the likelihood (l) needs to be evaluated by experts first, following the considerations recommended by Khorsandi and Aven (2017): What could cause the assumption to deviate? What are the key drivers of those causes? Has a similar deviation occurred in the past? What evidence is available for supporting the potential for a deviation?

Then, the value of B is determined based on the likelihood (l):

- a. $B = 1, \text{if } 0 \leq l \leq 20\%$
- b. $B = 2, \text{if } 20\% < l \leq 30\%$

c. $B = 3, \text{if } 30\% < l \leq 100\%$

Note that the belief of assumption deviation is discretized into three levels to facilitate a clear classification of different criticality levels of the assumption deviation risks through the decision tree models, which will be discussed in details in Sect. 2.2.9.

2.2.4. Evaluate the amount of believed deviation from the true value

The amount of believed deviation is evaluated as the relative distance between the assumed parameter value and the true value believed by experts, as expressed by Eq. (2). Similar to the belief in deviation, the believed deviation D is evaluated by experts and represents the experts' belief on how severe the deviation could be. The value assigned to D takes a positive sign (+) if the assumption is believed to deviate towards dangerous scenarios and a negative sign (−) if it is deviating towards safe scenarios:

$$D = \frac{p_t - p}{p} \quad (2)$$

where D is the amount of believed deviation, p_t is the parameter value believed true by the experts, and p is the parameter value as assumed in the analysis.

2.2.5. Evaluate the strength of knowledge

The assigned belief (likelihood) and amount of deviation are conditional on the background knowledge available, and on the individual expertise and points of view of the experts who made the assessment. Therefore, the strength of knowledge on which the assessment is based is highly relevant and is explicitly considered in both the original and extended framework. In this paper, we use the method proposed by Flage and Aven (2009) for evaluating the strength of knowledge. This approach is mainly based on the evaluation of four criteria: (i) reasonability and realism of assumptions; (ii) phenomenological understanding; (iii) availability of reliable data and information; (iv) agreements among peers. In addition, we take into account a fifth criteria, suggested by Khorsandi and Aven (2017): (v) the level of expertise and competence of the experts. A score of 1-3 is given for each criterion, corresponding to three levels, i.e., weak, moderate and strong, respectively.

A weighted average of the five criteria scores $k_i, i = 1, 2, \dots, 5$, is used to calculate the overall knowledge score SK :

$$SK = \sum_{i=1}^5 w_i \cdot k_i, \quad (3)$$

where w_i is the weight of criterion k_i . Obviously, the five criteria are not equally important in defining the strength of knowledge. To handle this, the Analytical Hierarchy Process (AHP) (Saaty, 2008) is used to determine the weights

of the strength of knowledge criteria. A good feature of the method is that it can be helpful in group decision-making (Saaty, 2008). Experts are asked to fill pairwise comparison matrixes that represent the relative importances of the five criteria in defining the knowledge. The eigenvector problem is, then, solved and the weights are found by normalizing the principal eigenvector. The calculated weights from the experts are, then, averaged to give the final weights shown in Table 1.

Table 1 Strength of knowledge criteria and their weights.

Attribute	Weight
Reasonability and realism of assumptions (k_1)	0.13
Availability of reliable data and information (k_2)	0.13
Phenomenological understanding (k_3)	0.42
Agreement among peers (k_4)	0.16
Level of expertise and competence of the experts (k_5)	0.16

The strength of knowledge denoted by K , is, then, calculated based on the value of SK :

- $K = 1$, if $1 \leq SK \leq 1.6$
- $K = 2$, if $1.6 < SK \leq 2.3$
- $K = 3$, if $SK > 2.3$

Note that the knowledge on assumption deviation is discretized into three levels to facilitate a clear classification of different criticality levels of the assumption deviation risks through the decision tree models, which will be discussed in details in Sect. 2.2.9.

2.2.6. Determine the context of decision

In the original work of Khorsandi and Aven (2017), only one context of decision making was considered, i.e., comparing a risk metric to a specific safety objective. In this sense, only assumptions deviations toward dangerous scenarios need to be considered. In the practice of risk management, however, we often need to compare alternatives in terms of their risks. In this case, all the deviation scenarios need to be considered, since a conservative assumption might lead to a higher risk metric, which again leads the decision maker to prefer other alternatives; in other words, it gives a “false alarm” of high risk. For more illustration, take the example in Figure 5. In this example, the decision maker is comparing two alternatives, Al_1 and Al_2 , and he/she prefers to choose the alternative with the lower risk. At a first glance, the decision maker would choose Al_1 as it has the lowest risk metric value (the blue solid line).

However, a second look shows that the value of R_2 (in the meshed filling) is lower than that of R_1 , when the true condition is used in the calculation rather than a conservative assumption.

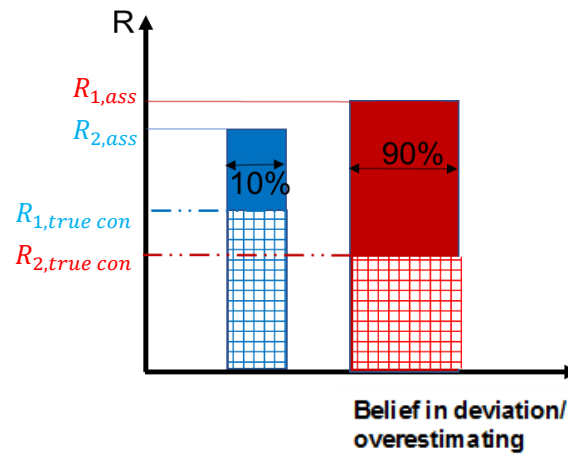


Figure 5 Comparing the risk related to two alternatives taking into account the risk metric value based on the assumption made and the true condition.

Hence, it is important to identify the context of decision making when implementing the extended framework. In this paper, two decision making contexts are distinguished, namely, comparing a risk metric to a safety objective (DM_1) and comparing two alternatives (DM_2).

2.2.7. Define the safety objective

The safety objective needs to be identified considering the given decision context, as shown in Figure 3. The safety objective represents a numerical value whose exceedance by the risk metric would lead to changes in the results of the risk-informed decision making. The safety objective is dependent on the context of the decision making. For the decision context DM_1 , the safety objective is identified as the threshold that the risk metric should not exceed. On the other hand, if the decision context is DM_2 , the assessor needs to choose the alternative with the lowest risk metric value. Therefore, the (higher) risk metric value of another alternative is defined as the safety objective under this decision making context.

2.2.8. Identify the margin of deviation

Next, the margin of deviation (μ) needs to be calculated. This margin represents the maximum tolerable assumption deviation before the risk-informed decision is changed. As shown in Figure 4, different assumptions might affect one or more model parameters, or, the other way around, a model parameter might be affected by one or more assumptions. In this paper, we calculate the margin of deviation one assumption at a time, to reduce the complexity of the analysis. Assume that the assumption of interest a_i affects model parameters p_1, p_2, \dots, p_m . Then,

we assume that the assumption deviation affects “similarly” the related parameters (p_1, p_2, \dots, p_m) to make the equation solvable. The assumption deviation can be modeled by:

$$\begin{cases} p'_1 = (1 + \mu)p_1 \\ p'_2 = (1 + \mu)p_2 \\ \vdots \\ p'_m = (1 + \mu)p_m \end{cases} \quad (4)$$

where p'_i , $i = 1, 2, \dots, m$, are the deviated model parameters and μ represents the amount of deviation in the model parameters (and assumed to be the same for all parameters affected by an assumption) due to the deviation in the assumption. It should be noted that in theory, the basic event probabilities can also change by different amounts, resulting in different values of μ for different basic events. Then, the deviated risk metric \hat{R} is calculated by

$$\hat{R} = f(p'_1, p'_2, \dots, p'_m, p_{m+1}, \dots, p_n) \quad (5)$$

The value of μ can be calculated by solving the following equation:

$$\underset{\mu}{arg} f((1 + \mu) \cdot p_1, (1 + \mu) \cdot p_2, \dots, (1 + \mu) \cdot p_m, p_{m+1}, \dots, p_n) = R_{th} \quad (6)$$

In Eq. (6), R_{th} is the safety objective defined in Sect. 2.2.7, i.e.:

$$R_{th} = \begin{cases} R_{lim}, & \text{if the decision context is } DM_1 \\ R_2, & \text{if the decision context is } DM_2 \end{cases} \quad (7)$$

where R_{lim} and R_2 represent the safety limit objective and the risk metric value of the alternative being compared, respectively.

2.2.9. Evaluate the overall criticality based on the decision flow diagrams

The criticality of an assumption deviation measures its influence on the risk-informed decision making and, hence, on the safety of the system. As defined in Sect. 2.1, the criticality of the assumption deviation depends on both the severity of the influence and the likelihood of the deviation. Four scenarios are distinguished to quantify the severity of the influence of the assumption deviation:

- a. failures in meeting the established objectives, i.e., the magnitude of deviation is larger than the deviation margin, leading to the exceedance of the safety limit;
- b. success in meeting the established objectives i.e., the magnitude of deviation is lower than the deviation margin, or the deviation is occurring towards lower amounts of risk due to conservatism in the assumption;
- c. Altering the different prioritization when comparing two or more alternatives, i.e., the risk metric based on unrealistic assumptions is higher or lower than what it would be based on the true conditions, leading to the mischoice among the different alternatives.

- d. Unchanging the prioritization when comparing two or more alternatives, i.e., the risk metric based on unrealistic assumptions is higher or lower than what it would be based on the true conditions, leading to misranking the different alternatives.

Considering the scenarios defined above and the likelihood of deviation, decision flow diagrams are built in Figure 6-8 for evaluating the criticality of assumption deviation risk. The decision flow diagrams are introduced to facilitate the “standardization” of the evaluation of the assumption deviation effects on risk assessment. It should be noted that in these figures, the difference between the margin of deviation μ and the amount of deviation D , denoted by $\Delta\mu$, is calculated and used to measure the safety margin for a given assumption deviation:

$$\Delta\mu = \mu - D \quad (8)$$

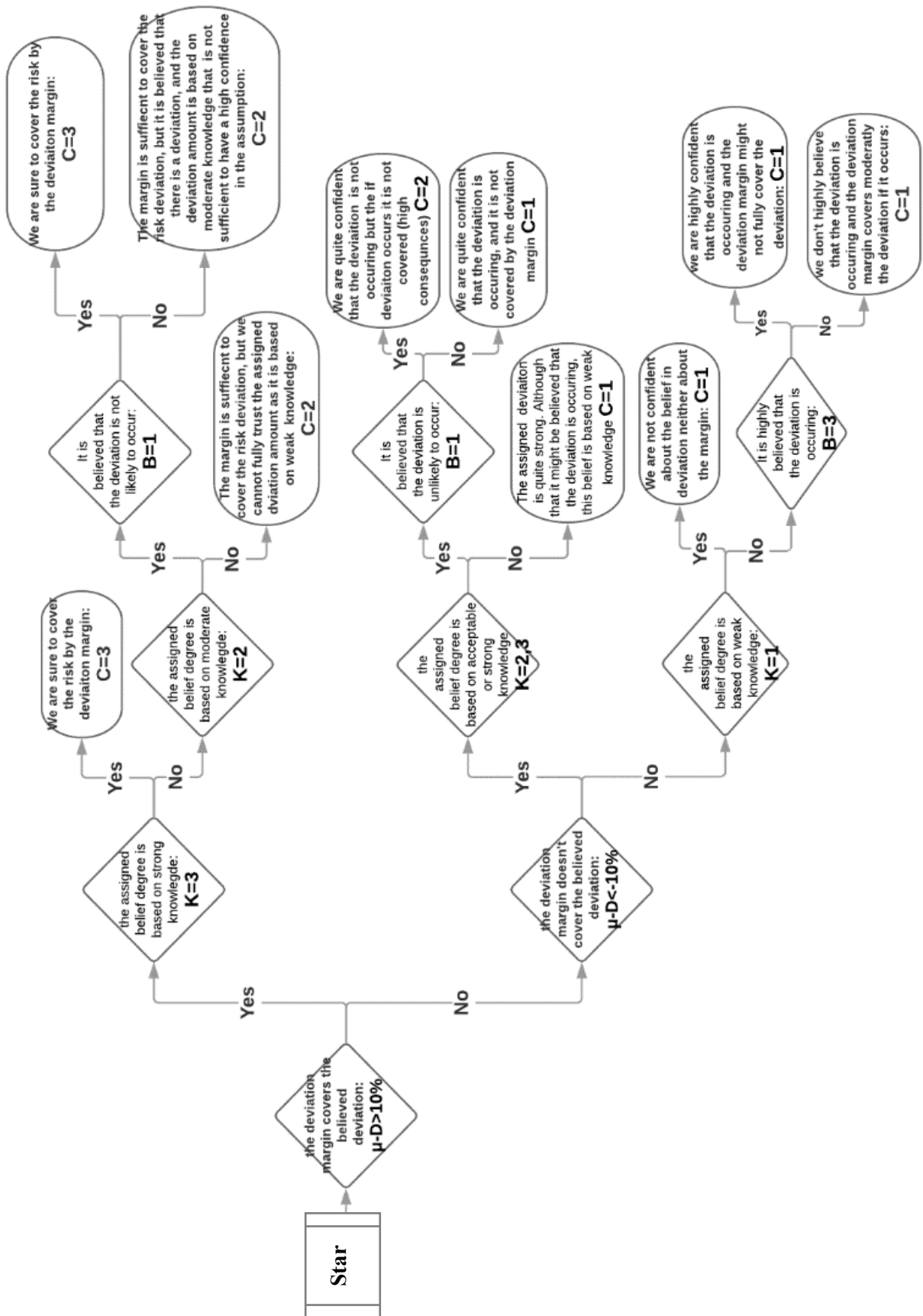


Figure 6 Criticality assessment decision flow diagram for decision context DM_1 and assumptions of types A_1 and A_2 .

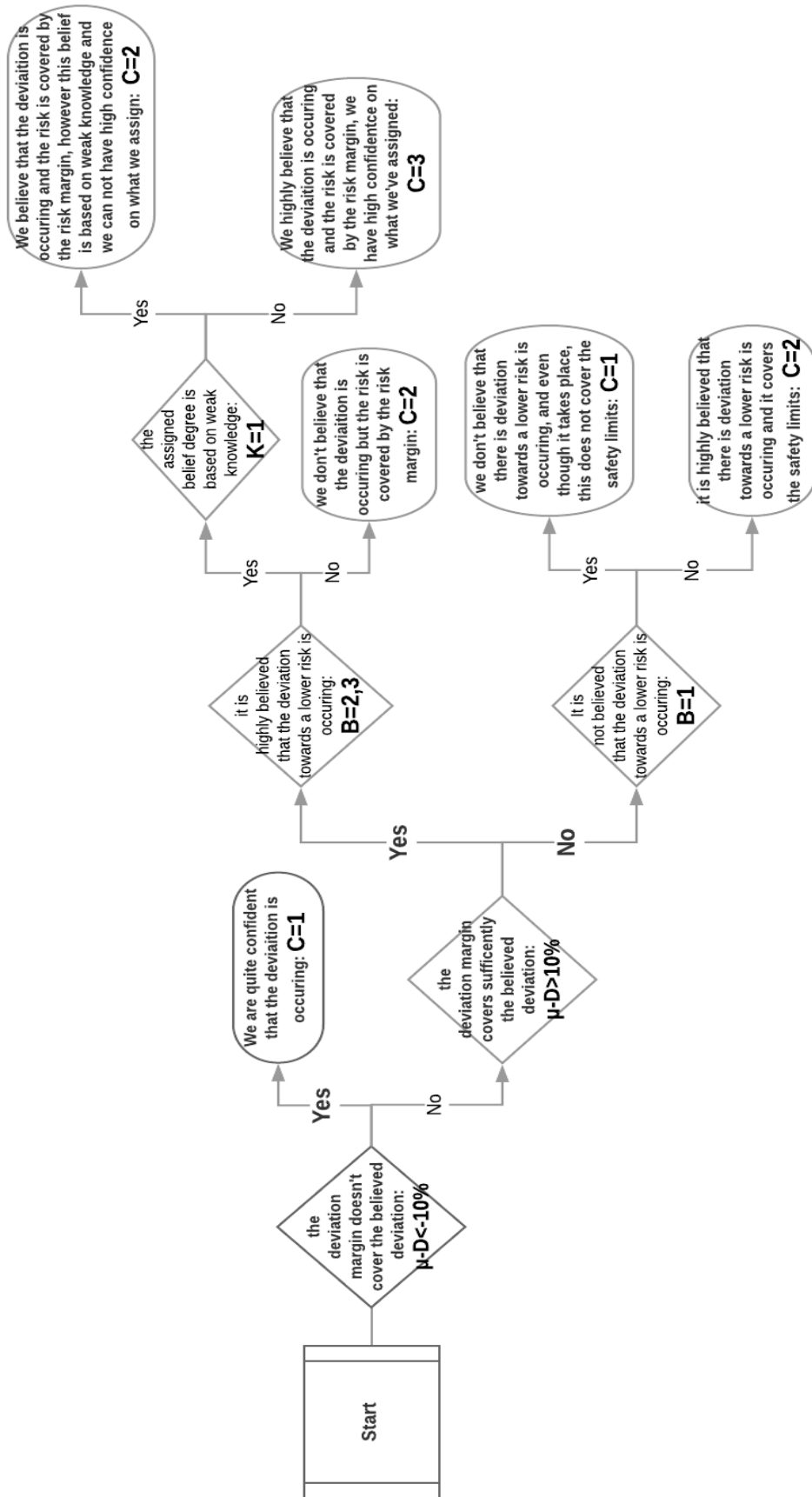


Figure 7 Criticality assessment decision flow diagram for decision context DM_1 and assumptions of type A_3 .

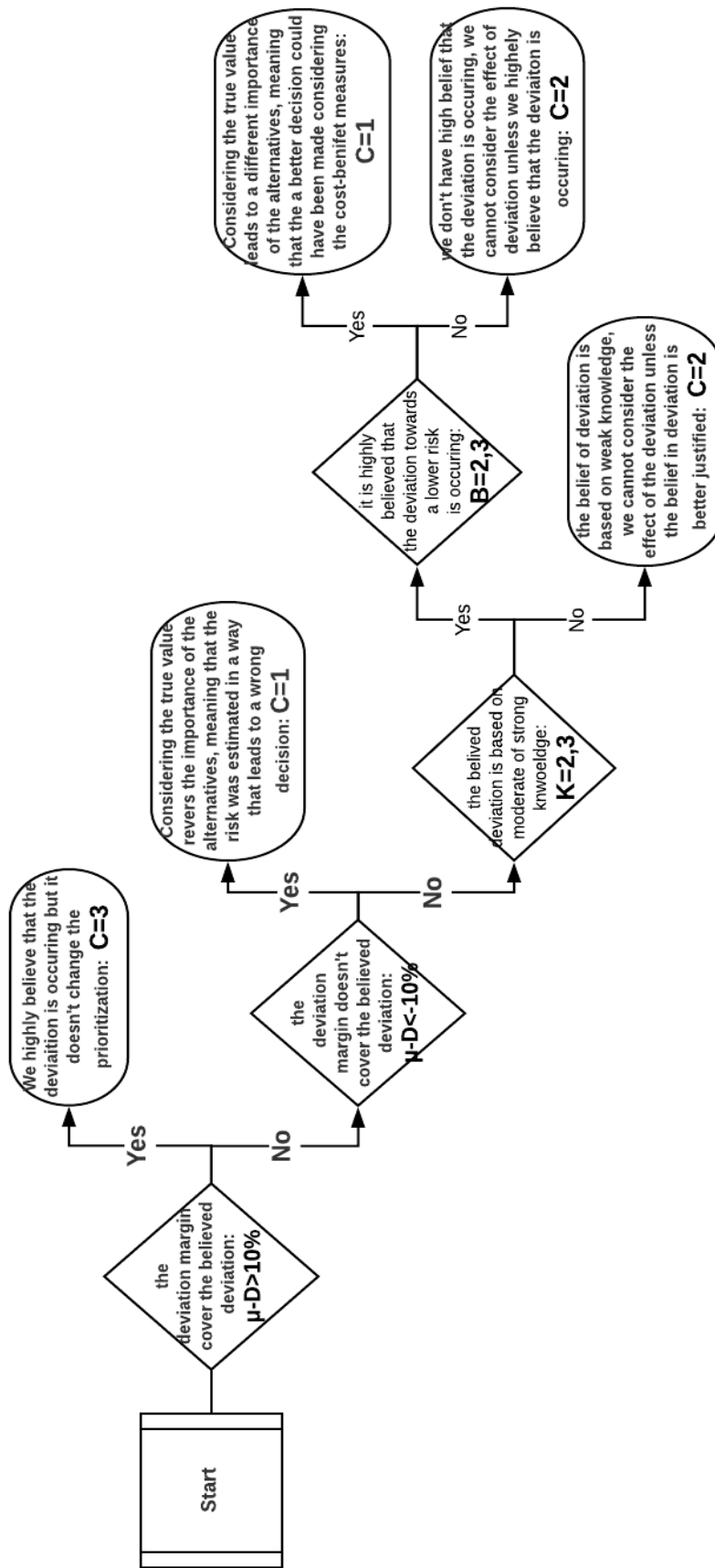


Figure 8 Criticality assessment decision flow diagram for decision context DM_2 and assumptions of types A_1 , A_2 and A_3 .

Following the steps in Sects. 2.2.1-2.2.9, the criticality C can be evaluated using the decision flow diagrams in Figures 6-8. Take the case in Table 2 as an illustrative example. In this example, the assessor assigns a 90% probability of deviation, meaning that he or she is almost sure that the assumption deviates from reality. The amount of the believed deviation is evaluated to be 20%. The two values are assigned based on strong knowledge, i.e., $K = 3$, which means that the assessment is judged to be credible to a certain degree and can be trusted. The difference between the deviation margin and the amount of the believed deviation is 40%. This logically means that we are sure to be under the safety limits even though the real condition deviates from the assumption. However, as the decision context in this example is DM_1 and the type of assumption is A_2 , the decision flow diagram in Figure 6 is chosen for evaluating C . It can be seen from Figure 6 that in this case, we have $C = 3$, meaning that the assumption can be trusted and that decisions can be made based on the current assumption, as the assumption deviation risk is judged to be low.

Table 2 An example of a classification of assumptions deviation risk.

Criteria	Assessment
Type of assumption (A_i)	Best judgment
Context of decision making (DM_i)	Comparing the risk metric to a risk limit
Likelihood of deviation (l)	90%
Amount of believed deviation (D)	20%
Strength of knowledge (K)	Strong
Margin of deviation (μ)	60%

3. Case study

In this section, we apply the developed framework on a case study of real PRA models for the external flooding hazard groups in an NPP. The PRA models were developed by Electricité de France (EDF). The needed data and information that supports the model development were found in the technical reports provided by EDF.

3.1. Description of the PRA model

PRA models are used for investigating undesired events and quantifying their likelihoods and consequences. Similar to all analytical models, PRA models are conditional on the models' assumptions (EPRI, 2015). The assumption made are mainly: (i) assumptions made in case of lack of information and understanding of some phenomena or risk-related aspects; (ii) assumptions made for reducing the complexity of the model and to make it operational (these assumptions are also called approximations in (EPRI, 2015)). The PRA model for external flooding is chosen because it is less mature compared to the PRA model of other hazard groups and involves many assumptions.

External flooding is a naturally induced hazard that might be caused due to different initiating events, such as river overflows, dam failures and snow melts (IAEA, 2003, IAEA, 2011). The PRA model developed by EDF is a combination of fault trees and event trees, evaluated under different scenarios, e.g., water levels and operation states. The model structure and the probabilities of basic events (BEs) are, in turn, related to specific assumptions made by experts. The original external flooding PRA model is of a large scale (i.e., it includes three operation states, thousands of BEs and several thousand Minimal Cut Sets (MCS), and a large number of assumptions). A reduced-order model has been constructed in Bani-Mustafa *et al.* (2018) to represent the original model with less complexity, i.e., less BEs and less MCSs. In this paper, we consider the reduced-order model in Bani-Mustafa *et al.* (2018) for assumption deviation risk assessment. In this reduced order model, only one operating state (Normal Shutdown with cooling using Steam Generator-NS/SG) that contributes to 86% of the risk metric value is considered. In this operating state, one scenario (water levels) whose risk contribution is 98.7% is considered. Given the operating state and scenario considered, 5 MCSs that contribute to 80.1% of the risk are considered. The corresponding MCSs and BEs of the reduced-order model are presented in Tables 3-4.

Table 3 Reduced-order model constituents (Bani-Mustafa *et al.* 2018).

Operating state	Scenarios	MCS
NS/SG	Water level A	MCS1={BE1, BE2, BE3}
		MCS2={BE2, BE3, BE4}
		MCS3={BE3, BE5, BE6, BE7, BE8}
		MCS4={BE2, BE3, BE7, BE9}
		MCS5={ BE2, BE3, BE6, BE10}

Table 4 Basic events included in the reduced-order model (Bani-Mustafa *et al.* 2018).

Symbol	Basic event
BE1	External flooding with water level A inducing a loss of offsite power
BE2	Loss of auxiliary feedwater system due to the failure to close the isolating valve
BE3	Loss of component cooling system because of clogging
BE4	Common cause leading to the failure of all pumps of the Auxiliary feedwater (AFW) system
BE5	Failure of the turbine of AFW system
BE6	Failure of the Diesel Generator A
BE7	Failure of the Diesel Generator B
BE8	Failure of the common diesel generator
BE9	Common cause leading to the failure of pumps 1 and 2 of AFW system
BE10	Common cause leading to the failure of pumps 2 and 3 of AFW system

Taking the rare-event approximation, the total risk metric R_{Red} of the reduced-order PRA model can be calculated by:

$$R_{Red} = \sum_{i=1}^{n_{O,Red}} \sum_{j=1}^{n_{S,Red,i}} \sum_{k=1}^{n_{MCS,Red,i,j}} \prod_{q \in MCS_{i,j,k}} P_{BE,q} \quad (9)$$

where $n_{O,Red}$ is the number of operation states in the reduced order model, $n_{S,Red,i}$ is the number of scenarios in the reduced-order model, $n_{MCS,Red,i,j}$ is the number of minimal cutsets in the reduced-order model, $P_{BE,q}$ are the probabilities of the basic events in the reduced-order model. As shown in Bani-Mustafa *et al.* (2018), using the reduced-order model allows reproducing approximately 68% of the total risk contribution.

3.2. Evaluation of assumption deviation risk

3.2.1. Identifying critical assumptions

The critical assumptions in the PRA model of external flooding are identified following the procedures in Sect. 2.2 and listed in Table 5. The assumption deviation risks for the assumptions in Table 5 need to be evaluated using the developed method in Sect 2. In the following, we illustrate in detail how to apply the developed framework on one conservative assumption, namely “the clogging accompanying some floods is unpredictable and unfilterable”. For the other assumptions, we directly give the classification results in Sect. 3.2.8.

Table 5 List of the assumptions related to the reduced-order model of the external flooding hazard group.

As_i	Description	Type	Affected basic event
As_1	It is assumed that failure to close the isolating valves for volumetric protection sealing-water proofing causes the total loss of Emergency Feed Water System (EFWS)	Conservative	BE2
As_2	If the floods occur, the clogging is certain ($P = 1$)	Best judgment	BE3
As_3	If the river flooding is accompanied with clogging, then, it is unpredictable and unfilterable	Conservative	BE3, BE4
As_4	Clogging leads to failure of Essential Services Water System (component cooling system) and therefore, the reactor containment spray system	Best judgment	BE3, BE4
As_5	It is assumed that probabilities of a given level of flood can be calculated by	Best judgment	BE1

	extrapolating the distributions based on observed data to the extreme water flowrate (i.e., flowrates that have never occurred) and that the probabilities of floods can be taken as mean values		
As_6	It is assumed that once the water reaches the bottom of an equipment, the equipment fails	Conservative	BE2-BE10
As_7	It is assumed that once the water level exceeds the height of the barriers, the water will enter and fill the building	Best judgment	BE2-BE10
As_8	It is assumed that unit 1 cannot get help from unit 2 and vice versa, or from the safeguard system shared between the two units	Conservative	BE8
As_9	It is assumed that the river flood can be predicted using statistical models	Optimistic	BE1
As_{10}	It assumed that once the river flood is predicted, the probability of failing to transit into the state of “emergency shutdown” (i.e., normal shutdown and cooling with steam generator, normal shutdown and cooling with residual heat removal system etc.) is the intrinsic failure probability that is considered in normal cases	Best judgment	BE1

3.2.2. Identification of model parameters affected by the assumption of interest

The model parameters in the PRA model are the probabilities of the basic events in the event tree. As the clogging can lead to the loss of component cooling system (CCS) or the loss of the pumps in the auxiliary feedwater system, the assumption As_3 is related to the two basic events BE3 and BE4, as presented in Table 5.

3.2.3. Assessment of the belief in deviation

Experts from EDF are invited to assess the belief in deviation. In this assumption, the probability that the clogging is not detected and filtered is 1 ($P = 1$), while in reality, the clogging is usually detectable and can be filtered, which means that the true value of this probability is less than 1 ($P < 1$), leading to a lower risk than the value calculated using the assumed model parameters. Therefore, the experts think that this assumption is very conservative, indicating that the assumption deviation might reduce the value of the risk metric.

Some observations can also help the expert to better understand the assumption and evaluate the belief in deviation, as shown in Table 6.

Table 6 Assessment of the belief in deviation for AS_3

Aspects	Assessment
What could cause the assumption to deviate?	<p>The amount of precipitation can usually be predicted. Hence, if the river flooding is caused by precipitation, then, it can be predicted.</p> <p>Unless it is due to barrier rupture, the river level usually increases gradually and can be seen and noticed easily.</p> <p>If there is heavy precipitation, the operators would pay more attention to the water filters on the river and clean the filters to make sure that the water intake is not clogged.</p>
What are the key drivers of those causes?	<p>The fact that the river level increases is a gradual process.</p> <p>The fact that the operators are able to clean the clogging if it occurs.</p>
Has a similar deviation occurred in the past?	Yes.
What evidence is available for supporting the potential for a deviation?	The feedback reports show that a clogging has occurred before and that operators were able to see it and manage it.

Based on the analysis illustrated in Table 6, the belief in deviation was assigned to be 70%. Therefore, we have

$$B = 3.$$

3.2.4. Evaluate the amount of believed deviation from the true value

Experts in EDF are asked to evaluate, based on their beliefs, the amount of assumption deviation from the true values. The experts have assigned the amount of deviation in percentage to be $D = -50\%$, meaning that the experts believe that the assumption is conservative and deviating towards a higher risk.

3.2.5. Evaluate the strength of knowledge

The strength of knowledge has been evaluated as indicated in Sect. 2.2.5. The strength of knowledge attributes are evaluated separately, as shown in Table 7.

Table 7 Strength of knowledge criteria and weights.

Attribute	Weight	Score
Reasonability and realism of assumptions (k_1)	0.13	1
Availability of reliable data and information (k_2)	0.13	2
Phenomenological understanding (k_3)	0.42	1
Agreement among peers (k_4)	0.16	1
Level of expertise and competence of the experts (k_5)	0.16	2

The overall knowledge score K is calculated using Eq. (3):

$$K = \sum_{i=1}^5 w_i \cdot k_i = 1.29$$

Then, based on the criteria defined in Sect. 2.2.5, we have $K = 1$.

3.2.6. Determine the context of decision making and define the safety objective

The context of the decision making in this case study is to compare a risk metric to a safety limit. The risk limit for core meltdown varies between 1×10^{-5} and 1×10^{-4} (Knochenhauer and Holmberg, 2012). As the flooding events are usually site-specific (IAEA, 2009), the contribution of the external flooding hazard group to core meltdown also varies from one NPP to another. Moreover, we consider only a part of the external flooding PRA model in this case study (through the reduced-order model). Accordingly, for illustration purposes, we artificially set the safety limit of the considered PRA model to be $R_{lim} = 1.6 \times 10^{-8}$.

3.2.7. Identify the margin of deviation

As the assumption As_3 affects the basic events BE_3 , BE_4 , the vector of basic events' probabilities related to the assumption are $P_m = (p_{BE_3}, p_{BE_4})$. Accordingly, the deviated risk function can be expressed using Eq. (5):

$$R' = R_{th} = R_{lim} = f(p_1, p_2, p'_{BE_3}, p'_{BE_4}, p_5, \dots, p_{10}) = f(p_1, p_2, (1 + \mu) \cdot p_3, (1 + \mu) \cdot p_4, p_5 \dots p_{10})$$

The solver in Microsoft Excel is used to solve Eq. (6), with $R_{lim} = 1.603 \times 10^{-8}$. The resulted margin of deviation is $\mu_{As_3} = 26.40\%$. The margins of deviation for the remaining assumptions are calculated in a similar way, as presented in Table 8 next in Sect. 3.2.8.

3.2.8. Evaluate the overall criticality based on the decision flow diagram

As illustrated in Sect. 2, the overall criticality of assumptions deviation is assigned based on the decision flow diagrams in Figure 6-8. For the assumption of interest (AS_3), the belief (likelihood) in the deviation is assigned to be 70% (level 3). The difference between the deviation margin and the amount of believed deviation is 76.40%. The strength of knowledge is assessed to be $K = 1$. For an acceptance-criteria decision-context, this means that we believe that we are under the safety limit, and the deviation is not considered critical and can be accepted. On the other hand, our belief is based on weak knowledge, which makes it less credible. Following the decision flow diagram in Figure 6, the criticality of this assumption is $C = 2$. Accordingly, the assumption is not very critical and listed in the “waiting list”, which means that it is accepted unless there are other criteria and information on other assumptions deviations that change the evaluation.

The same steps are repeated for each assumption. The scores and the evaluation corresponding to each criterion for each assumption are presented in Table 8 together with their final criticality scores.

Table 8 Assumption-deviation criticality and criticality criteria assessment

A_i	Type	BEs	$l_i : B_i$	D_i	μ_i	$\Delta\mu_i$	K_i	C_i
1	Conservative	BE2	95%:3	-90%	∞	∞	1	2
2	Best judgment	BE3	30%:2	90%	35.11%	-54.89%	2	1
3	Conservative	BE3, BE4	70%:3	-90%	26.40%	116.40%	1	2
4	Best judgment	BE3, BE4	5%:1	5%	26.40%	21.40%	3	3
5	Best judgment	BE1	50%:3	50%	24.22%	-25.78%	3	1
6	Conservative	BE2-BE10	90%:3	-70%	20.38%	90.38%	1	2
7	Best judgment	BE2-BE10	40%:3	30%	20.38%	-9.62%	2	1
8	Conservative	BE8	20%:1	-30%	869.95%	899.95%	1	2
9	Optimistic	BE1	40%:3	30%	24.22%	-5.78%	2	1
10	Best judgment	BE1	5%:1	5%	24.22%	19.22%	3	3

As shown in Table 8, the different assumptions have three levels of criticality i.e., 1; 2; 3 (very critical; not very critical; not critical). The corresponding actions that need to be taken by decision makers and analysts are respectively:

- (i) $C = 3$: The deviation is very likely to happen. Besides, the assumption deviation has severe influence on the decision making and might lead to exceedance of the safety limit. Further analysis and justification of the assumption is required. This kind of assumptions decreases greatly the safety margin of the NPP. Therefore, it should be treated carefully.
- (ii) $C = 2$: The assumption can be trusted to support decision making if the risks of the deviation from other assumptions are all not critical ($C = 3$). Further analysis and justification of the assumption is needed only when other assumptions are also in this state. This kind of assumptions does not decrease the safety margin of the NPP if the other assumptions are of the same type or less critical.
- (iii) $C = 1$: An assumption deviation is unlikely to happen or, if it happens, it does not affect the decision making nor the safety of the NPP. The assumption can be trusted and decisions can be made based on the current assumption. This assumption does not impact the safety margin of the NPP.

As shown from the example above, the assumptions deviations might be inevitable. Since they might significantly affect the results of QRA, the decision makers and analysts should pay attention to their criticality. In the NPP industry in particular, some deviations might be very critical and lead to catastrophic consequences.

4. Discussion and conclusions

In this paper, we have extended the approach of Khorsandi and Aven (2017) for evaluating assumptions deviations in probabilistic/quantitative risk assessments. The extended framework covers a new context of decision making very relevant in practice, namely, that of comparing alternatives (rather than comparing a single alternative against a safety objective) and an additional type of assumptions, namely, conservative assumptions (rather than just best judgment and optimistic types of assumptions). An integrated metric, the criticality of assumption deviation, is defined and evaluated based on the extended framework through the use of decision flow diagrams. The developed framework includes a detailed procedure for evaluating the criticality of assumptions deviations using six criteria: (i) the type of assumption; (ii) the context of decision making; (iii) the belief in the deviation from reality; (iv) the amount of deviation from reality; (v) the margin of deviation; (vi) the strength of the knowledge supporting the assumption made. The criticality is, then, evaluated based on decision flow diagrams. These diagrams depict a logical flow based on the aforementioned criteria to facilitate a standardized classification of the criticality of assumptions deviation based on its impact on safety. Alongside the decision diagrams, a detailed applicational procedure for evaluating the criteria in the framework has been presented to improve its applicability. Also, a reduced-order model

has been introduced for reducing the scale of the analysis by focusing the analysis on the most important assumptions.

The developed framework is applied to a case study of a PRA model of the external flooding hazard group of an NPP. The implementation of the framework has shown its feasibility and its ability to cover different types of assumptions, and to provide a more transparent and complete evaluation of the assumption deviation. In particular, the case study shows that, in addition to the optimistic and best judgment assumptions that might be a source of safety margins reduction, there are assumptions that are very likely to be greatly overestimating the true values (conservative). Such a kind of assumptions could be very problematic in case the decision maker is to choose between alternatives, as they lead to an overestimation of the risk, which could result in a sub-optimal decision.

The use of decision flow diagrams has both pros and cons. The pros are that these diagrams facilitate a standardized assumption deviation risk assessment, increasing both the transparency and efficiency of the assessment. These are desirable attributes in case of peer review of the assessment and considering the large number of assumptions typically involved in PRAs. A con of such diagrams are that they give a “mechanical” assessment procedure where the assessment is based on strict rules rather than the use of overall judgements. Another possible limitation of the current research that need to be addressed in the future is that it analyzes the deviation risk for one assumption at a time and, thus, fails to take into account the deviation risk for several assumptions simultaneously.

Acknowledgements

The work on this article was performed in part while R. Flage was visiting CentraleSupélec in the period February-March 2018. He would like to acknowledge the financial support from CentraleSupélec, as well as his co-author Professor E. Zio for making his stay possible, and memorable.

References

- Aven, T. (2013). Practical implications of the new risk perspectives. *Reliability Engineering & System Safety*, 115, 136-145.
- Andrews, J. D. and Moss, T. R. (2002) *Reliability and risk assessment*. Wiley-Blackwell.
- Apostolakis, G. E. (2004) ‘How useful is quantitative risk assessment?’, *Risk Analysis: An International Journal*. Wiley Online Library, 24(3), pp. 515–520.
- Aven, T. (2011) *Quantitative risk assessment: the scientific platform*. Cambridge University Press.
- Aven, T. and Zio, E. (2013) ‘Model output uncertainty in risk assessment’, *International Journal of Performability Engineering*, 9(5), pp. 475–486.
- Bani-mustafa, T. *et al.* (2018) ‘Strength of Knowledge Assessment for Risk Informed Decision Making’, in *Esrel*.

Trondheim.

Bedford, T. and Cooke, R. (2001) *Probabilistic risk analysis: foundations and methods*. Cambridge University Press.

Berner, C. and Flage, R. (2016) 'Strengthening quantitative risk assessments by systematic treatment of uncertain assumptions', *Reliability Engineering and System Safety*, 151, pp. 46–59. doi: 10.1016/j.ress.2015.10.009.

Berner, C. L. and Flage, R. (2017) 'Creating risk management strategies based on uncertain assumptions and aspects from assumption-based planning', *Reliability Engineering & System Safety*. Elsevier, 167, pp. 10–19.

Bjerga, T., Aven, T. and Zio, E. (2014) 'An illustration of the use of an approach for treating model uncertainties in risk assessment', *Reliability Engineering and System Safety*, 125, pp. 46–53. doi: 10.1016/j.ress.2014.01.014.

Boone, I. *et al.* (2010) 'A method to evaluate the quality of assumptions in quantitative microbial risk assessment', *Journal of Risk Research*, 13(3), pp. 337–352. doi: 10.1080/13669870903564574.

Cox Jr, L. A. (2012) *Risk analysis foundations, models, and methods*. Springer Science & Business Media.

Eiser, J. *et al.* (2012) 'Risk interpretation and action: A conceptual framework for responses to natural hazards', *International Journal of Disaster Risk Reduction*, 1(1), pp. 5–16. doi: 10.1016/j.ijdr.2012.05.002.

EPRI (2006) *Guideline for the Treatment of Uncertainty in Risk-Informed Applications: Applications Guide, 2006*. Palo Alto, California: 1013491.

EPRI (2015) *An Approach to Risk Aggregation for Risk-Informed Decision-Making*. Palo Alto, California.

Fullwood, R. R. and Hall, R. E. (1988) 'Probabilistic risk assessment in the nuclear power industry'.

Goerlandt, F. and Montewka, J. (2014) 'Expressing and communicating uncertainty and bias in relation to Quantitative Risk Analysis', *Safety and Reliability: Methodology and Applications*, 2(13), pp. 1691–1699. doi: 10.1201/b17399-230.

Henley, E. J. and Kumamoto, H. (1981) *Reliability engineering and risk assessment*. Prentice-Hall Englewood Cliffs (NJ).

IAEA (2003) *External Events Excluding Earthquakes in the Design of Nuclear Power Plants*.

IAEA (2009) *Meteorological and Hydrological Hazards in Site Evaluation for Nuclear Installations*.

IAEA (2011) 'IAEA-Publication8635'.

Kaplan, S. and Garrick, B. J. (1981) 'On the quantitative definition of risk', *Risk analysis*. Wiley Online Library, 1(1), pp. 11–27.

Khorsandi, J. and Aven, T. (2017) 'Incorporating assumption deviation risk in quantitative risk assessments: A semi-quantitative approach', *Reliability Engineering & System Safety*, 163, pp. 22–32. doi:

<https://doi.org/10.1016/j.ress.2017.01.018>.

- Kloprogge, P., Van der Sluijs, J. P. and Petersen, A. C. (2011) 'A method for the analysis of assumptions in model-based environmental assessments', *Environmental Modelling and Software*. Elsevier Ltd, 26(3), pp. 289–301. doi: 10.1016/j.envsoft.2009.06.009.
- Knochenhauer, M. and Holmberg, J. E. (2012) 'Guidance for the definition and application of probabilistic safety criteria', *Proceedings of PSAM 10 International Probabilistic Safety Assessment & Management*.
- Netjasov, F. and Janic, M. (2008) 'A review of research on risk and safety modelling in civil aviation', *Journal of Air Transport Management*. Elsevier, 14(4), pp. 213–220.
- NRC (1975) *Reactor safety study: An assessment of accident risks in US commercial nuclear power plants*. Hein.
- NRC (2010) *Reactor Coolant System and Connected Systems, NUREG-0800 Standard Review Plan for the Review of Safety Analysis Reports for Nuclear Power Plants*. Washington: NRC.
- Saaty, T. L. (2008) 'Decision making with the analytic hierarchy process', *International Journal of Services Sciences*, 1(1), p. 83. doi: 10.1504/IJSSCI.2008.017590.
- Stewart, M. and Melchers, R. E. (1997) *Probabilistic risk assessment of engineering systems*. Springer.
- Vesely, W. E. and Apostolakis, G. E. (1999) 'Developments in risk-informed decision-making for nuclear power plants', *Reliability Engineering & System Safety*. Elsevier, 63(3), pp. 223–224.
- Vinnem, J.-E. (2014a) *Offshore Risk Assessment, Vol 1.: Principles, Modelling and Applications of QRA Studies*. Springer.
- Vinnem, J.-E. (2014b) *Offshore Risk Assessment, Vol 2.: Principles, Modelling and Applications of QRA Studies*. Springer.
- Vose, D. (2008) *Risk analysis: a quantitative guide*. John Wiley & Sons.
- Zio, E. (2007) *An introduction to the basics of reliability and risk analysis*. World scientific.
- Zio, E. (2009) *Computational methods for reliability and risk analysis*. World Scientific Publishing Company.
- Suter II, G. W. (2007). *Ecological risk assessment*. Second Edition. Taylor & Francis.