

# A Framework for Multi-Hazards Risk Aggregation Considering Risk Model Maturity Levels

Tasneem Bani-Mustafa<sup>1</sup>, Zhiguo Zeng<sup>1</sup>, Enrico Zio<sup>1,2</sup>, Dominique Vasseur<sup>3</sup>

<sup>1</sup> 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

<sup>2</sup> Energy Department, Politecnico di Milano, Via Ponzio 34/3, Milan, 20133, Italy

<sup>3</sup> 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  
e-mail: tasneem-adeeb.bani-mustafa@ecp.fr

**Abstract**—Multi-Hazards Risk Aggregation (MHRA) aggregates risk over different risk contributors and provides a final risk index that permits the comparison with safety guidelines to support Decision Making (DM). Maturity of the risk analysis are always different for different risk contributors, resulting in different degrees of confidence on the associated results. The difference in maturity, and the sources of heterogeneity that cause such differences, should be taken into account for supporting a reliable and accurate representation of risk in respect of DM.

In this paper, we first propose a hierarchical framework to evaluate the level of maturity of risk assessments in respect of DM. The framework consists of different attributes (uncertainty, conservatism, knowledge and sensitivity) and sub-attributes (availability of data, consistency of data, data reliability, experience, and value ladenness) that are believed to determine the level of maturity of a risk assessment. Evaluation guidelines are pre-defined for facilitating the attributes scoring process. Then, the maturity level is integrated in MHRA by developing a two-dimensional risk aggregation method. Finally, a numerical case study for the MHRA of a Nuclear Power Plant (NPP) is carried out to show the applicability and the plausibility of the methods.

**Keywords**—risk assessment; probabilistic safety assessment; level of maturity; risk-informed decision making; multi-hazards risk aggregation

## I. INTRODUCTION

Probability Safety Assessment (PSA) has been widely applied in various industries, e.g., nuclear, aerospace, defense, etc. Usually, different hazard groups are involved in a PSA (e.g., PSA of nuclear power plants usually involves hazard groups like fire, internal flooding, etc.). To make risk-informed decisions based on the results of PSA, Multi-Hazard Risk Aggregation (MHRA) is required: all relevant information on risk from different contributors is combined, arriving at an integrated risk index [1]. Usually, risk-informed decisions are made by comparing the integrated risk index (e.g., core damage frequency, large early release frequency, risk increase, etc.) to safety goals and quantitative acceptance criteria.

Currently, most MHRAs are conducted by summing over the individual risk indexes for different hazard groups [1].

For example, in current PSAs for Nuclear Power Plants (NPPs) in France, an overall risk index is computed by summing over the risk indexes of hazard groups like internal events, fire, etc. The overall risk index is, then, compared to safety goals and acceptance guidelines for Risk-Informed Decision Making (RIDM) [1].

A main criticism for the summation-based MHRA method is that it ignores the heterogeneities in the trust we have on the knowledge we have on the different hazard groups. Take again the PSA of nuclear power plants as an example. Among the hazard groups, the PSA model has been developed for internal events for many years, while relatively recently, the PSA for hazard groups like fire has started to be investigated [1]. Therefore, we have more trust on the PSA for internal events than for fire. The different levels of trust, which result from the difference in knowledge that supports the risk assessments, must be taken into account as they affect the risk-informed decisions based on the results of risk analyses [2].

In this paper, we develop a new MHRA method that considers the heterogeneities of trust on the risk assessment results from different hazard groups. The trust on a risk model is evaluated based on the concept of maturity. Maturity of a PSA is defined in this paper as the degree to which a PSA is correctly implemented in a way to reflect the available knowledge. The rest of this paper is organized as follows. In Sect. 2, we present a hierarchical framework for assessing the maturity of PSA. Then, in Sect. 3, we develop an MHRA method that considers the maturity of the PSAs for different hazard groups. Section 4 applies the developed methods on a numerical case study. Finally, in Sect 5, the paper is concluded with a discussion on potential future works.

## II. A HIERARCHICAL FRAMEWORK FOR PSA MATURITY ASSESSMENT

Various works attempt to assess the trust and validity of PSAs, e.g., [1-8]. Different elements have been used for maturity assessment, e.g., uncertainty, state of knowledge, importance of assumptions, difference in level of details and analysis of different risk contributors, difference in level of conservatism over different hazard groups, accuracy and precision in the estimation of the values of the model's

parameters, degree of sensitivity, etc. In this paper, we develop a hierarchical framework for maturity assessment based on these elements, as shown in Figure 1.

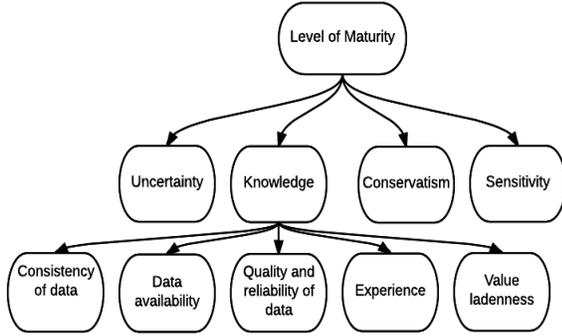


Figure 1. Level of maturity framework

As shown in Figure 1, four criteria (attributes) are defined for maturity assessment:

- Uncertainty is defined as the imperfection of knowledge of the real value of a variable or its variability [3]. The uncertainty is an important source of differences between the actual values and the values predicted by a model [4]. This means that it can directly reflect the level of maturity of the analysis and it should be addressed when evaluating the level of maturity of a risk analysis.
- Knowledge is the yield of the combination of data and information with experience and judgment that support decision-making [5]. It is natural that building a credible and mature analysis of risk is conditional on the robustness and strength of knowledge. In other words, knowledge is a direct indication of the level of maturity of risk analysis and it should be presented in the maturity model.
- Conservatism in risk management is related to the need of overestimating risk rather than underestimating, that arises from different considerations and perspectives such as concerns regarding the lack of knowledge about the nature and magnitude of the hazard [6]. In general, conservatism in PSA leads to overestimating the risks. The degree of conservatism should match the requirement of the PSA or otherwise, it might reduce the maturity level and mislead the decision maker.
- Sensitivity refers to how the risk model output changes when the input parameters or the background knowledge change. In this paper, we mainly focus on the sensitivity of the assumptions made in the PSA, which indicates the importance of the assumptions made and the uncertainty associated with them [7].

The attribute knowledge is further broken down into five sub-attributes to make its evaluation, as shown in Figure 1. For each criterion and sub-criterion defined in Figure 1, a

semi-quantitative evaluation is carried out by assigning to it a relative score from 1 to 5, based on a set of pre-defined scoring criteria. For details of scoring criteria, please refer to the Appendix.

Then, the scores for the criteria and sub-criteria are aggregated for maturity assessment. In this paper, as a first attempt, all the criteria and sub-criteria are considered equally important and the maturity level is calculated as a weighted average of the scores of the criteria:

$$m_h = \frac{1}{N} \sum_{j=1}^N \frac{1}{n} \sum_{i=1}^n (Sc)_{i,j}, \quad (1)$$

where  $m_h$  is the level of maturity for hazard group  $h$ ,  $Sc_{ij}$  is the score for the  $j$ -th sub-criterion in the  $i$ -th evaluation criteria,  $N$  is the total number of criteria and  $n$  is the number of sub-criteria related to the  $i$ th evaluation criterion.

### III. RISK AGGREGATION CONSIDERING MATURITY LEVELS

In this paper, we adopt the perspectives of [2] that when characterizing risk, not only the probability index estimated by PSA, but also the knowledge that supports the PSA should be taken into account. Hence, in this paper, we use a tuple  $(R_i, M_i)$  to quantify the risk associated with hazard group  $i$ , where  $R_i$  is the risk index estimated from the PSA model and  $M_i$  is the maturity of the PSA, evaluated based on the method presented in Sect. 2.

A two-stage aggregation method is, then, developed for MHRA considering maturities of hazard groups. Suppose we have  $n$  hazard groups with the risk tuple  $(R_i, M_i)$ ,  $i = 1, 2, \dots, n$ . The overall risk tuple can, then, be computed in two steps:

Step 1: Aggregation of risk indexes. Risk indexes are aggregated following the summation rule:

$$R = \sum_{i=1}^n R_i, \quad (2)$$

where  $R$  is the risk index after considering all the hazard groups. The physical meaning of  $R$  is the aggregated risk index, when we have complete confidence on each of the hazard group.

Step 2: Determine the maturity of the aggregated risk assessment:

$$M = \min_{1 \leq i \leq n} M_i, \quad (3)$$

where  $M$  is the maturity level of the PSA considering all the hazard groups. It should be noted that by using the formula, we assume that the least mature hazard group determines the maturity of the entire assessment process, i.e., it follows the weakest-link model.

The aggregated risk, denoted by the tuple  $(R, M)$ , can, then, be used to support risk-informed decision making. Suppose we are considering the risk of a specific event. Instead of directly comparing  $R$  to the acceptance threshold, the maturity level should also be considered: when maturity level is low, a larger safety margin is required; while when maturity level is high, a risk close to its threshold value might be accepted. The relationship between maturity level and the required safety margin should be determined, based on the severity of the consequence of the event.

#### IV. NUMERICAL CASE STUDY

In this case study, we are conducting the risk analysis of a NPP, where the quantification of risk is made by a PSA model. For RIDM, a summation is performed over different hazard contributors to permit the comparison of the yielded risk metrics to acceptance criteria. The risk contributors are classified in the ASME/ANS standards into different hazard groups: (i) internal events; (ii) internal floods; (iii) internal fires; (iv) seismic events; (v) high winds; (vi) external floods; (vii) other hazards [8]. In this example, only two main important hazard groups, i.e., Internal events (I) and External flooding (Ex), are considered for illustrative purposes.

PSA of these two hazard groups are conducted based on a combination of fault tree and event tree models. The resulted risk indexes are:

$$R_I = 3 * \frac{10^{-8}}{\text{reactor.year}} \quad \text{and} \quad R_{Ex} = 2 * \frac{10^{-5}}{\text{reactor.year}}$$

The two hazard groups are different in nature and the knowledge and experience regarding them is quite different.

Applying the maturity framework of Figure 1, a score has been given to each attribute and sub-attribute, with the help of the guidelines indicated in Appendix 1. Suppose we have, for the internal events:

$$Sc_1(I) = 4, Sc_2(I) = 4, Sc_4(I) = 3, Sc_{3,1}(I) = 5, Sc_{3,2}(I) = 3, Sc_{3,3}(I) = 4, Sc_{3,4}(I) = 4, Sc_{3,5}(I) = 4, Sc_{3,6}(I) = 4, Sc_{3,7}(I) = 3, Sc_{3,8}(I) = 4 \text{ and } Sc_{3,9}(I) = 3$$

$$\text{and for external flooding: } Sc_1(Ex) = 2, Sc_2(Ex) = 3, Sc_4(Ex) = 3, Sc_{3,1}(Ex) = 2, Sc_{3,2}(Ex) = 3, Sc_{3,3}(Ex) = 3, Sc_{3,4}(Ex) = 3, Sc_{3,5}(Ex) = 4, Sc_{3,6}(Ex) = 3, Sc_{3,7}(Ex) = 2, Sc_{3,8}(Ex) = 2 \text{ and } Sc_{3,9}(Ex) = 3.$$

where  $Sc_1, Sc_2, Sc_3$  and  $Sc_4$  correspond to level of uncertainty, conservatism, knowledge and sensitivity respectively.  $Sc_{3,1} \dots Sc_{3,9}$  are illustrated in Table A4.1.

Then, applying Eq. (1) on the previous scores yields  $m_h(I) = 3.7$  and  $m_h(Ex) = 2.7$ . A two-dimensional representation of the level risk estimate and its level of maturity is presented in Figure 2.

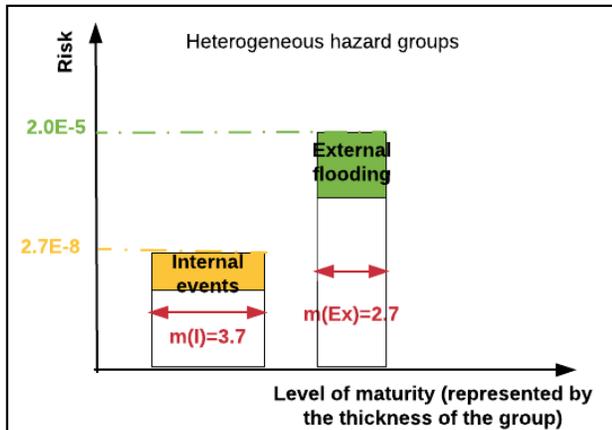


Figure 2. Representation of Hazard Groups Levels of Risk and Maturity

$$MHRA \text{ is, then, conducted, based on Eqs. (2) and (3):}$$

$$R = R_I + R_{Ex} = 3 \times 10^{-8} + 2 \times 10^{-5} = 2.003 \times 10^{-5}$$

$$M = \min_{1 \leq i \leq 2} (3.7, 2.7) = 2.7$$

Hence, the aggregated risk of the NPP can be represented as  $(R|m) = (2.003 \times 10^{-5}|2.7)$ .

In the previous analysis, although the risk due to external flooding is higher than for internal events, it should be noted that this value is based on a PSA with lower maturity.

#### V. CONCLUSION AND DISCUSSION

MHRA is very important for RIDM. Nevertheless, it is usually achieved using a simple arithmetic summation, which is not mathematically consistent nor physically meaningful due to the heterogeneous nature and level of maturity on the risk of the different risk contributors.

In this paper, we have proposed a framework that evaluates the level of maturity of risk analysis in support to DM. The model consists of different attributes that are believed to affect the maturity of risk analysis such as: uncertainty, conservatism etc. The maturity level is integrated in MHRA by developing a two-dimensional risk aggregation method. The method has been applied to a numerical example in order to show its applicability.

#### APPENDIX I EVALUATION OF SENSITIVITY

In this paper based on the literature review of several papers (e.g., [9]; [10]) and for simplicity, we suggest to use a one-at-a-time method index that measures the average of relative change in the dependent (output) variable  $Y(x_i)$  by altering one input ( $x$ ) as shown in 2.

$$S = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y(x_{i+1}) - Y(x_i)}{Y(x_{i+1})} \right|, \quad (A.1)$$

where  $x_i$  is the input parameter, n is the number of times that the analyst would apply the sensitivity measures by altering one input by an estimated suitable value (e.g.,  $\pm 20$ ).

In case of weak knowledge, the sensitivity analysis might be applied on input variables with ranges that are far from the true values which would be misleading. To overcome this problem, it suggested to alter the input parameters by a value equals to a standard deviation  $\pm SD$ .

In this kind of analysis converging from (0) indicates that the model is not sensitive, while the more the results diverges from (0) the more sensitive the model is. It might be that the interpretation would vary a lot and it would be non-meaningful to compare the sensitivity measures directly without trying to convert it into more meaningful scores that indicates in the light of the problem. A sensitivity score (1-5) is assigned for the sensitivity index relying on the degree that the index converge or diverge from 0 as illustrated in Table I.

TABLE I. SCORES REPRESENTATION OF THE SENSITIVITY MEASURE

Interval	Score	Score in the light of maturity
$S: \leq 0.5$	1	5
$S: 0.5-0.667$	2	4
$S: 0.667-0.8$	3	3
$S: 0.8-0.9$	4	2
$S: \geq 0.9$	5	1

APPENDIX II EVALUATION OF UNCERTAINTY

In this paper, we use the 5-level scaling system developed by Spiegelhalter and Riesch for uncertainty evaluation:

- Level 1 of uncertainty (level 5 of certainty): This level of uncertainty manifests itself when the model and the parameters are known, and the analysis predicts a certain outcome with a probability P , (e.g. the uncertainty about the outcome in most traditional mathematical and philosophical problems of probability theory) [3].
- Level 2 of uncertainty (level 4 of certainty): where the model is known but its parameters are not. If the parameters are known then the model would predict an outcome with probability P and exhibit an uncertainty of level one. In other words, this type of uncertainty arises due to lack of empirical information, (e.g. large break in primary circuit of a NPP that has never happened) [3].
- Level 3 of uncertainty (level 3 of certainty): Uncertainty about the model’s ability to reflect reality. This type of uncertainty is due to the model structure itself and the model computer implementation (Defining Uncertainty A Conceptual Basis for Uncertainty Management in Model-Based Decision Support, 2003). Moreover, out of the models that we can use, how likely each competing model is able to reflect reality [3].

- Level 4 of uncertainty (level 2 of certainty): uncertainty about acknowledged inadequacies and the assumptions that are implicitly made (non-modeled uncertainty). In fact, it arises from data extrapolations, limitation in the computations, and any aspects that we are aware that they have been omitted [3].
- Level 5 of uncertainty (level 1 of certainty): Uncertainty about unknown deficiencies and inadequacies and it is corresponding to unforeseen events, non-modeled and non-modable risks. It is the uncertainty that we are not aware of or as it was referred to by Donald Rumsfeld “unknown unknowns” (2002). In fact, this type of uncertainty are usually acknowledged by brainstorming of that possibilities that we can imagine and think of, or by the introduction of what so called ‘fudge factors’[3].

APPENDIX III EVALUATION OF KNOWLEDGE

In Table II, we tried to expand the knowledge characteristics suggested by [7] and adapt them to fit the risk assessment of NPPs. It is suggested to evaluate the level of knowledge relying on previously prepared questionnaire that has multiple answers each with specific score (1; 2; 3; and even number are left for intermediate cases). The scores are then averaged over all the characteristics and a final score is given for the knowledge.

TABLE II. LEVLE OF KNOWLEDGES' ATTRIBUTES EVALUATION GUIDELINES

	Score	1		3	5	
Data availability	Amount of data/field data ( $Sc_{3,1}$ )	No data or the data are so limited and (can extracted only from the same type of NPPs)		The data are available and can be extracted from any other NPP	The data are Available in abundance (can be extracted easily from so many sources and places worldwide)	
Data consistency	Source of data ( $Sc_{3,2}$ )	The data are extracted from other sources that is not related directly to the technology (not the exact same type of component)		Other NPPs of the same type and technology	Field data from the same power plant, and related to the same type of components	
Quality and reliability of data	Quality of Data ( $Sc_{3,3}$ )	Assumption of data	Data are calculated using statistical models	Data are both assumed and calculated using computer physical and mathematical models	Data are extracted using computer mathematical and physical models	The data are measured precisely and accurately, and then modeled
	Quality of assumptions ( $Sc_{3,4}$ )	Represents strong simplifications		Represents moderate simplifications	Represents reasonable simplifications	
Experience	Phenomenological understanding ( $Sc_{3,5}$ )	The phenomena involved are not well understood		The phenomena involved are Understood but not completely	The phenomena involved are very well understood	
	Experience and knowledge regarding the hazard group ( $Sc_{3,6}$ )	No experience at all		Experienced such an event in other industries	This event is quite common and we have a wide experience in	
	Availability of models ( $Sc_{3,7}$ )	Models are non-existent or known to give poor predictions.		The models used are believed to give predictions with moderate accuracy	The models used are known to give predictions with the required accuracy	
Value ladenness of the analysts	Agreement among peers ( $Sc_{3,8}$ )	There is strong disagreement among experts		There is slight agreement among experts	There is broad agreement among experts	
	Expert years in experience in the field and performance measure ( $Sc_{3,9}$ )	has quite short experience in risk assessment of NPPs		It is his specialty and he practiced through training courses regarding the same type of NPPs	Expert in this domain (long experience)	

## APPENDIX IV EVALUATION OF CONSERVATISM

In Figure 3 (a, b and c), based on the work of Aven (i.e., taking into account the meaning of conservatism, relating the conservatism to the strength of knowledge in which the probabilities are based and evaluating the usefulness of conservatism in the light decision-making) [11], we tried to extrapolate a more tangible semi-quantitative representation of the conservatism with respect to the level of maturity.

Type of estimate	Purpose	The conservative assumptions	Conservatism effect and evaluation with respect to the level of maturity
Best estimate	Comparison to a reference acceptance value	Higher than acceptance reference	Best estimate is higher than acceptance (4)
		Lower than acceptance reference	Best estimate is lower than acceptance (might be misinforming in terms of cost-benefit measures) (3)
	Comparing alternatives	Agrees with best estimate	Do not affect the decision (4)
		Disagrees with best estimates	Increases the confidence in the best estimate (3) The conservatism is misinforming in terms of cost-benefit risk reduction (2)

(a)

Type of estimate	Purpose	The conservative assumptions	Conservatism effect and evaluation with respect to the level of maturity
True Value (High confidence $P \geq 90\%$ ); based on weak knowledge	Comparison to a reference acceptance value	The conservative metric is higher than acceptance reference	True value is higher than acceptance (4)
		The conservative metric is lower than acceptance reference	True value is lower than acceptance might be misinforming in terms of cost-benefit measures) (2-3)
	Comparing alternatives	Agrees with true value	Do not affect the decision (4)
		Disagrees with true value	Increases the confidence in the true value (3-4) The conservatism is misinforming in terms of cost-benefit risk reduction (2)

(b)

Type of estimate	Purpose	The conservative assumptions is	Conservatism effect and evaluation with respect to the level of maturity
True Value (High confidence $P \geq 90\%$ ); based on strong knowledge	Comparison to a reference acceptance value	The conservative metric is higher than acceptance reference	True value is higher than acceptance (4-5)
		The conservative metric is lower than acceptance reference	True value is lower than acceptance might be misinforming in terms of cost-benefit measures) (2)
	Comparing alternatives	Agrees with true value	Do not affect the decision (5)
		Disagrees with true value	Do not affect a lot the decision (4-5) The conservatism is misinforming (1-2) in terms of cost-benefit analysis

(c)

Figure 3. Evaluation of the conservatism in the light of level of maturity (a, b, c: (conservatism VS Best estimate, True value/weak knowledge, True value/strong knowledge).

## REFERENCES

- [1] EPRI, "An Approach to Risk Aggregation for Risk-Informed Decision-Making," Palo Alto, California, 2015.
- [2] T. Aven, "Practical implications of the new risk perspectives," *Reliab. Eng. Syst. Saf.*, vol. 115, pp. 136–145, 2013.
- [3] H. Riesch, "Uncertainty," in *Essentials of Risk Theory*, no. 2009, 2013, pp. 29–57.
- [4] T. Bjerga, T. AVEN, AND E. ZIO, "An illustration of the use of an approach for treating model uncertainties in risk assessment," *Reliab. Eng. Syst. Saf.*, vol. 125, pp. 46–53, 2014.
- [5] T. Aven, "A conceptual framework for linking risk and the elements of the data-information-knowledge-wisdom (DIKW) hierarchy," *Reliab. Eng. Syst. Saf.*, vol. 111, pp. 30–36, 2013.
- [6] R. M. PERHAC JR, "Does risk aversion make a case for conservatism," *Risk*, vol. 7, p. 297, 1996.
- [7] FLAGE AND T. AVEN, "Expressing and communicating uncertainty in relation to quantitative risk analysis," *Reliab. Risk Anal. Theory Appl.*, vol. 2, no. 13, pp. 9–18, 2009.
- [8] ASME, "Addenda to ASME / ANS RA-S – 2008 Standard for Level 1 / Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications," 2009.
- [9] D. M. HAMBY, "A review of techniques for parameter sensitivity analysis of environmental models," *Environ. Monit. Assess.*, vol. 32, no. 2, pp. 135–154, 1994.
- [10] I. M. Sobol, "Sensitivity estimates for nonlinear mathematical models," *Math. Model. Comput. Exp.*, vol. 1, no. 4, pp. 407–414, 1993.
- [11] T. Aven, "On the use of conservatism in risk assessments," *Reliab. Eng. Syst. Saf.*, vol. 146, pp. 33–38, 2016.