Addressing Multi-Hazards Risk Aggregation for Nuclear Power Plants

Tasneem Bani-Mustafa

31/08/2015

Addressing Multi-Hazards Risk Aggregation for Nuclear Power Plants

Level of Maturity Model Building

BANI-MUSTAFA Tasneem
Supervisor: VASSEUR Dominique
EDF

Table of Content

List of Figures

List of Tables

1. Context

The safety of French nuclear reactors is based essentially on a deterministic approach. Probabilistic safety assessment (PSA) supplement the conventional deterministic analyses. A PSA is a systematic and comprehensive methodology to evaluate risks associated with a complex engineering technology entity (such as airlines or a nuclear power plant). It provides an overall view of safety including both equipment and operators behavior.

The development of PSA consist of:

- Identifying accidental scenarios leading to undesired consequences
- Assessing these scenarios in terms of frequency

In the year 1995 NRC has recommended in its final policy statement to increase the use of PSA in nuclear regulatory activities to the extent supported by the state of the art [1]. The current PSA models for plants in operation are built to estimate the risk for unit. The initiating events in PSA models may be classified in different hazard groups:

- So-called internal events.
- Internal hazards (fire, internal flooding).
- External hazards (external flood, earthquake, strong winds, etc.).

PSA model was developed over many years for internal events, but specific models have recently been developed for each hazard. The level of conservatism of the models is based on the origin of the initiating events. In particular, for external hazards, due to the lack of data (testing, physical models, etc.) conservative assumptions can be made regarding the impact of the hazards on the installation. The fact that the development of PSA it became a point of interest comes from the importance of these aspects in risk-informed decision making process.

"There is too much focus on the letter P[1] in the PSA". It has been said once that in PSA we tend to focus only on the probability and the consequences, and we ignore a lot of other important factors that plays a big role in the total risk and risk analysis process. Actually a lot of researches proved that there are a lot of other factors beside the risk (probability × consequences) that might affect the result of total risk aggregation. Through the operation of US nuclear power plant, fire has been considered as a great contributor of the total risk, which might be due to the importance of the fire risk or/and due to that it is characterized as immature and less realistic compared to some other initiating events; such as the internal events. As it is indicated in one of NRC's technical reports that was published earlier this year (2015) about the maturity of fire analysis, the fire analysis might be conservative compared to other hazards [2]. It is illustrated using the radar charts in ***Figure 2*** that represents CDF (core damage frequency) [3].

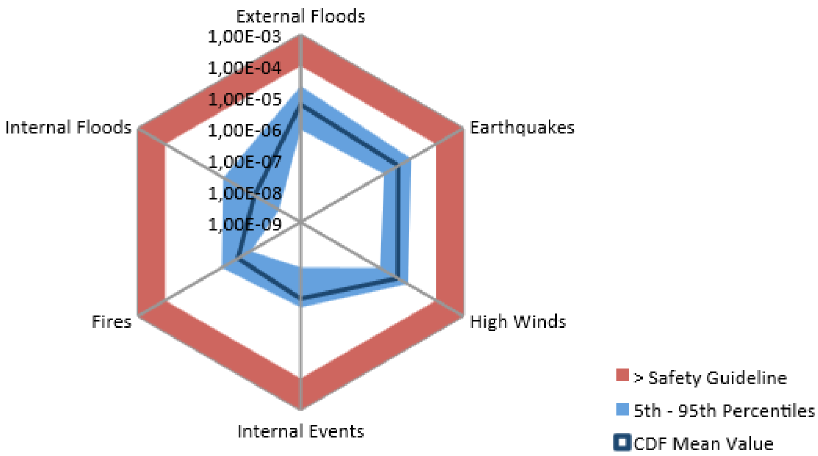

Figure 1 Example of radar chart representing CDF distribution of different hazard groups

In the other hand there are many other key aspects that it is believed to influence the risk informed decision making process, such as the interpretation of uncertainty

[1] Here P stands for probability

analysis, and the sensitivity of those analysis, as well as the other factors related to the decision maker himself see ***Figure 2***. Actually understand these aspects and their impact on PSA (probabilistic safety analysis), and providing a pragmatic approach of addressing these aspects is a critical point in development of the risk-informed decision making process [4], [5]. Similarly, for the evaluation of the frequency of these hazard, studied at extreme levels of intensity, it is often difficult to establish a result in which we can have a great confidence.

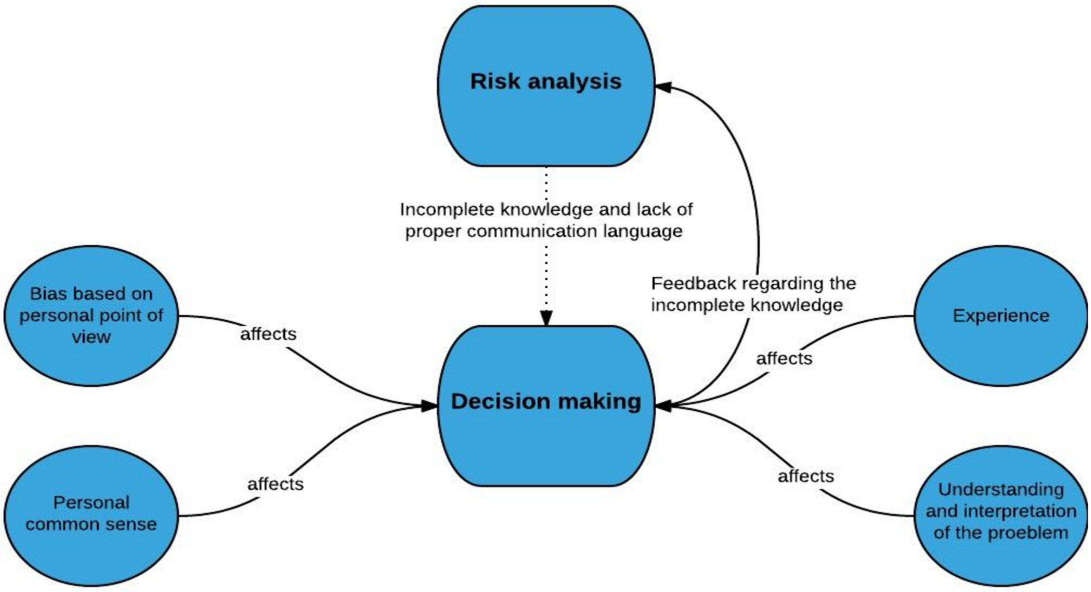

Figure 2 Risk informed decision making process, factors influencing decision making, and RIDM weaknesses

2. Definition of the Problem

In order to make decision we need to have a global comprehensive view of the risk, for example when we have to assess the benefits of a specific modification of the plant (ex: after Fukushima's accident design modifications as shown in ***Figure 3***). The risk might be evaluated taking into account the different origins and natures of the hazard groups; the following figure illustrates the different contributions of the different hazard groups before and after the modifications.

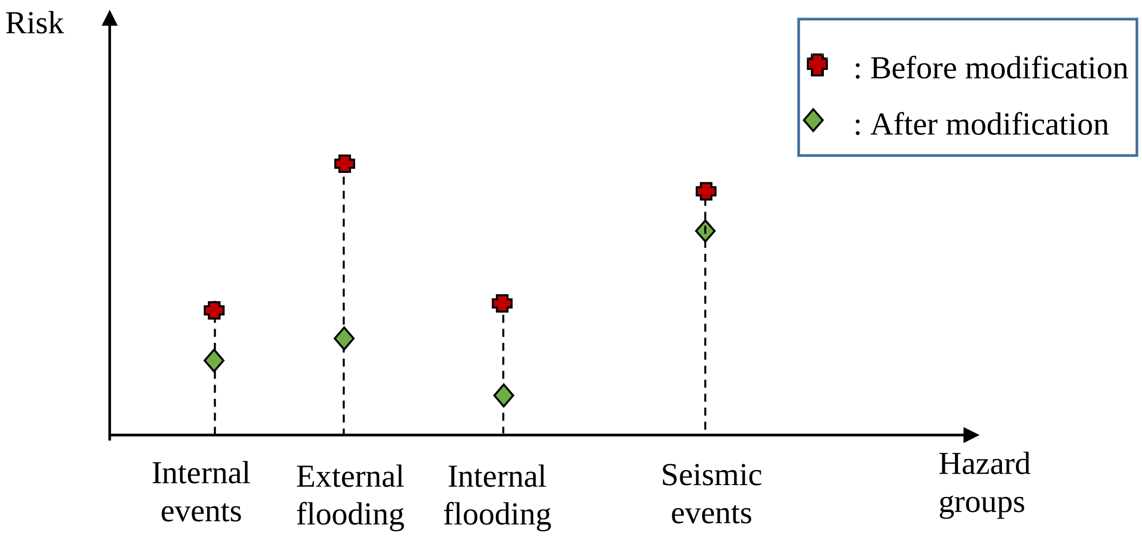

Figure 3: Risk contributions from different hazard groups before and after modifications

Risk Informed Decision making (RDIM) was defined by NASA in risk informed decision making handbook [6], as being "a deliberative process that uses a diverse set of performance measures, along with other considerations, to inform decision making. The RIDM process acknowledges the role that human judgment plays in decisions, and that technical information cannot be the sole basis for decision making. This is not only because of inevitable gaps in the technical information, but also because decision making is an inherently subjective, values-based enterprise. In the face of complex decision making involving multiple competing objectives, the cumulative wisdom provided by experienced personnel is essential for integrating technical and nontechnical factors to produce sound decisions."
In ***Figure 4*** which was taken from (NASA Risk-Informed Decision MakingHnadbook, 2010) the part deliberation is referred to as the communication gate between different parties and stages including the risk analysis and the decision making. Actually it is necessary to ensure the completeness, integrity, and comprehension of the information that is needed to be delivered and interpreted into a decision [6]. Aven also referred to the risk analysis as a decision supporting tool in which by, the decision maker should be able to balance the cost and the risk [7].
We agree that the risk analysis is part and parcel of decision making process, and that the analyst should understand the context of decision making, but still at the current stage we believe that the way we are using the risk analysis result in decision making is not enough, due to the different types of knowledge of those two processes, where one of them might be on the level of the implicit knowledge and the other is on the level of explicit knowledge that cannot be directly connected. So in order to make the right decision we need to know the risk, understand it, and of course to be able to acknowledge it in order to help the decision maker to comprehend it [8].

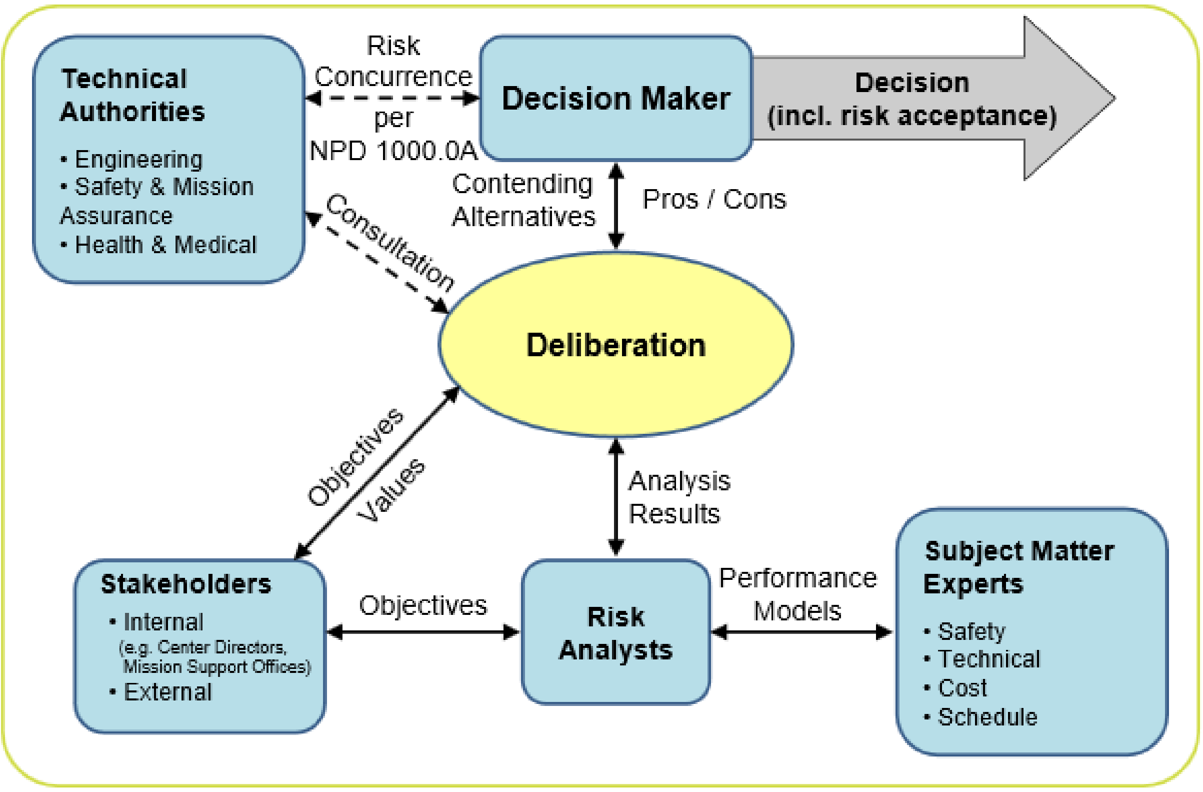

Figure 4: Functional Roles and Information Flow in RIDM (Notional)

In a philosophic way we can say that both decision maker and analysts speak a slightly different technical languages, because the decision maker decision is based more on the experience, sense, and wisdom even though that he uses the result of the analyst to make his decision. In the other hand the analysts rely more on data, facts, and numbers, which create a kind of technical gap between both of the analyst and the decision maker. It might be a real misleading issue for taking the right decision. The question is: are we able to translate the implicit knowledge into an explicit one, in order to come out with a new simplified technical language that is spoken by both parties?

In risk-informed decision making, risk aggregation is required to give an overall characterization of the risk, by combining different information on the risk from different contributors. The traditional approach of MHRA (multi hazard risk aggregation), which consists in simply summing risks from different hazard groups (risk contributors) as illustrated in ***Figure 5***, in order to come out with a final metric that should be used in order to evaluate the acceptance of the risk. This is not mathematically consistent nor physically meaningful, because of highly heterogeneous levels of parameter uncertainties, uncertainties for initiating event frequencies, modeling details and approximations, conservatisms in PSA models, and etc. Another problem arises from the fact that the classical approach does not address the issue of the potential interactions between risk factors that might increase the overall risk, and that some of PRA's methods are more conservative than the others, as well as that there are different natures of risks. The normal summation of the risk is going to give a mathematical result, but it won't give a physical meaning which might be sometimes meaningless and misleading [5], [9].

As it is explained before, this issue becomes a real concern when the results of PRA is to be used in risk informed decision making, due to the different levels of PSA's maturity [5], [9].

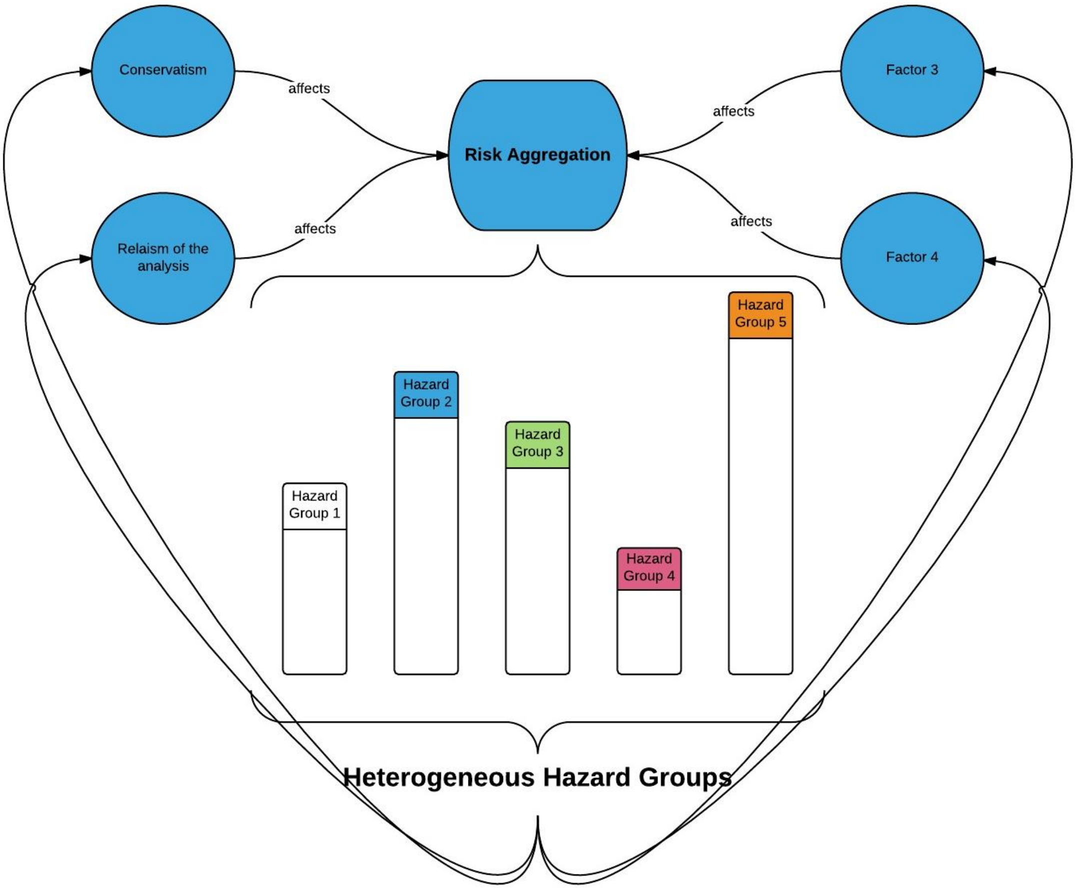

Figure 5 Risk aggregation from different hazard groups.

3. Existing Methods and Drawbacks

The aggregation of risks is defined in the EPRI document ([9], EPRI Technical Update, 2014) as a process that combines the information on the risk from different contributors in order to characterize risk and inform a decision. Given that in most of the "risk Informed" applications, it is necessary to compare a particular metric (e.g core damage frequency, large early release frequency, risk increase, etc.) to a threshold value or goal (e.g. RG 1.174 [2]), it is really tempting to do this aggregation through a "simple summation" of the different contributions to the metric studied. This summation may however present difficulties in terms of interpretation of the results, given the current limitations of PSA including hazards. The EPRI document (an Approach to Risk Aggregation for Risk-Informed Decision-Making) firstly explains the reasons why a direct summation can be dangerous in terms of decision making, and proposes an approach to overcome these difficulties within the RG.1174 context [2].

Actually, in an application informed risk decisions should be made by integrating several components. ***Figure 6*** below summarizes the aspects to be taken into account. This decision making context is taken from the RG 1.174 [2].

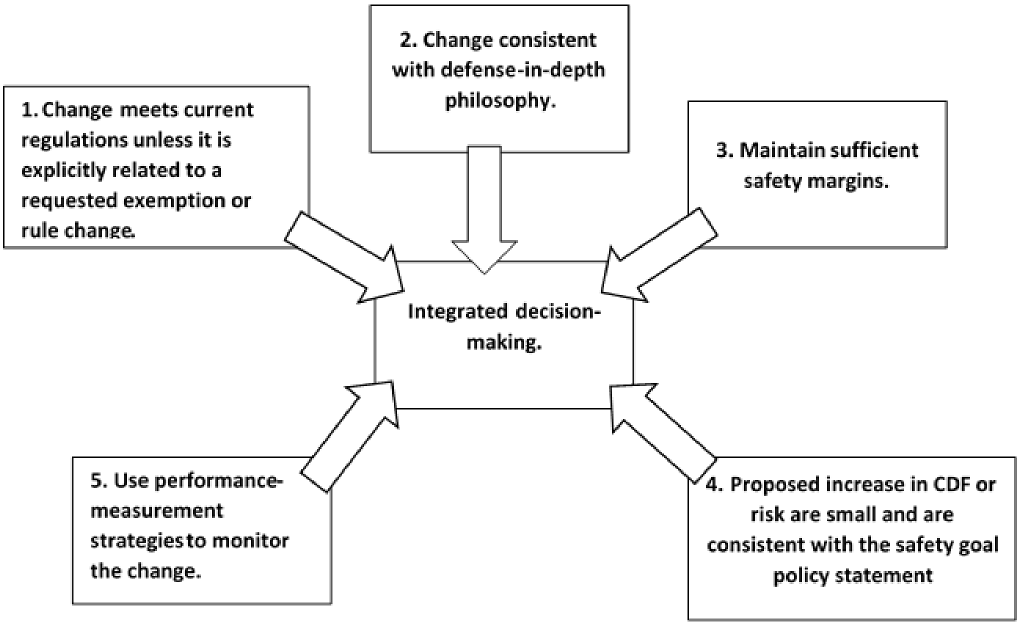

Figure 6: Risk Informed decision making process within RG 1.174 framework

However it is noted that the INSAG 25 of IAEA [3], offers the same type of approach. The use of quantitative result of PSA is clearly presented but is only one factor that enlightens the decision. However, it must include the uncertainty analysis that affect quantitative results ([9], EPRI Technical Update, 2014).

The approach proposed by EPRI can be summed up as in the following flowcharts *(**Figure 7**, **Figure 9**, and **Figure 8**).*

Task 2 aims to identify the main contributors to the baseline risk. Then, the analyst "disaggregates" the base model to establish how much credibility or confidence there is in the assessment. Important contributors should be reviewed in order to clearly understand the implications of departures from realism included in the PRA model. ***Figure 9*** outlines the basic process. ***Figure 8*** shows how to apply these tasks in the context of a risk-informed application that requires the evaluation of a risk metric, such as CDF or ΔCDF, by aggregating the results from PRAs for a number of different hazard groups.

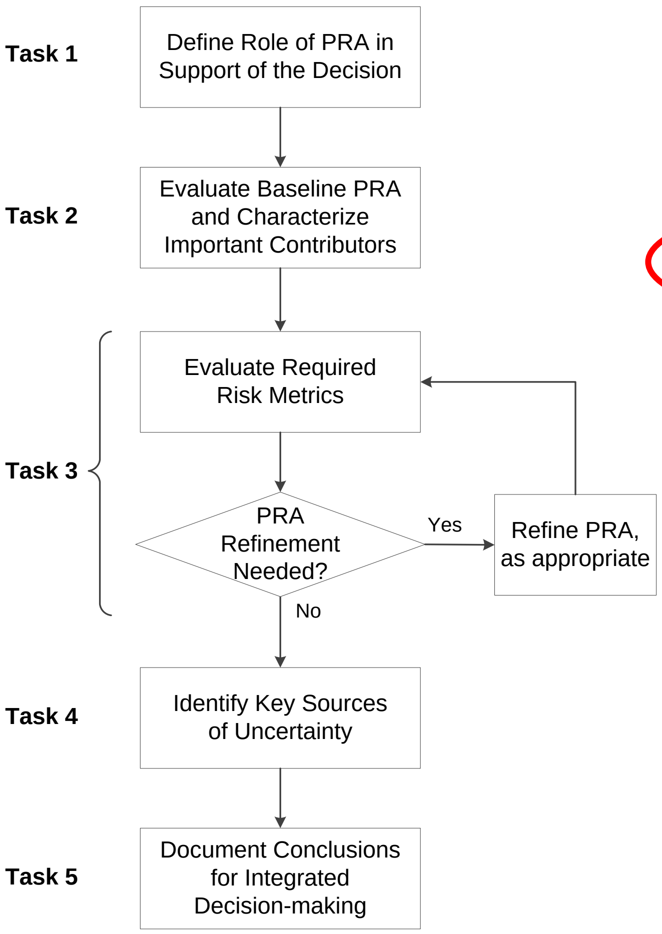

Figure 7 Proposed general approach

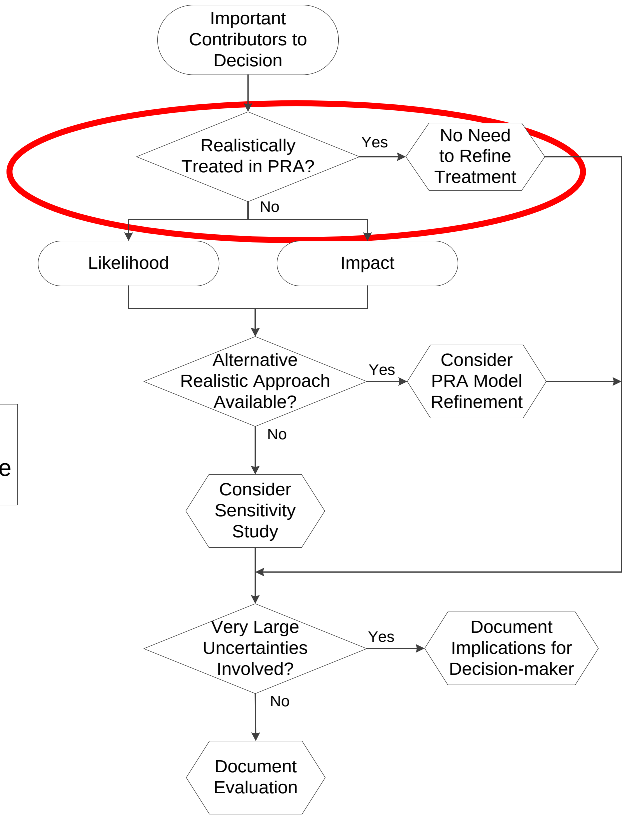

Figure 9 Process to assess realism of the PSA model

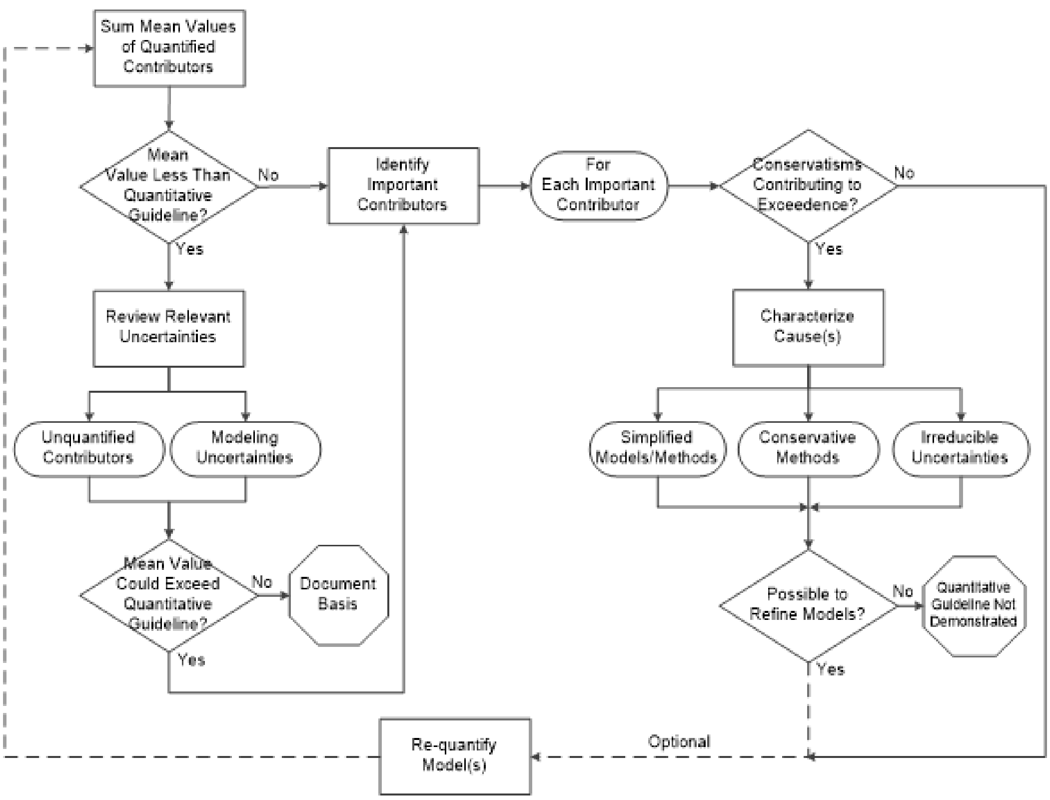

Figure 8: Process to aggregate

The EPRI document ([9], EPRI Technical Update, 2014) gives a good industrial view of the problem and propose an approach that fits very well to the "Risk Informed" philosophy. However, this document does not provide guidance on how the level of realism of a PSA model can be measured (see the marked area in the ***Figure 9***). Moreover, the presentation of the results of the assessment seems to be proposed without any reference to the decision making process.

An illustration of possible tools to be used to implement the process presented in ***Figure 8***, is given in the document (Addressing multi-hazards risk aggregation for nuclear power plants). This document proposes a new consistent approach to the quantifiable aspects of MHRA with a focus on relative rather than absolute risk metrics, using response surfaces based on arbitrary polynomial chaos expansion in combination with radar chart visualization of overall risk and associated uncertainties. Using the response surface, we can identify major contributors to overall plant risk, both on individual and aggregate bases, as well as cliff edge regions. While using arbitrary polynomial chaos (aPC). This method provides the mathematical tool to enable the investigations of MHRA, and it is capable of handling models with a very large number of input parameters (aPC feature), and any form of probability distribution, only knowing low-order statistical moments of these distributions. In other language, this method relies on quantifying the response surface rather than the complex original model itself.

Radar charts provide the analysis and communication tool to represent the multi-dimensional elements of the optimization process. Using the radar chart visualization tool, contributors of many different natures to selected metrics, are readily compared to regulatory safety guidelines. Ultimately, plant PSA models, response surfaces and radar charts can be combined into an iterative process (see ***Figure 10***), to support the process related to aggregate risk-informed decision making.

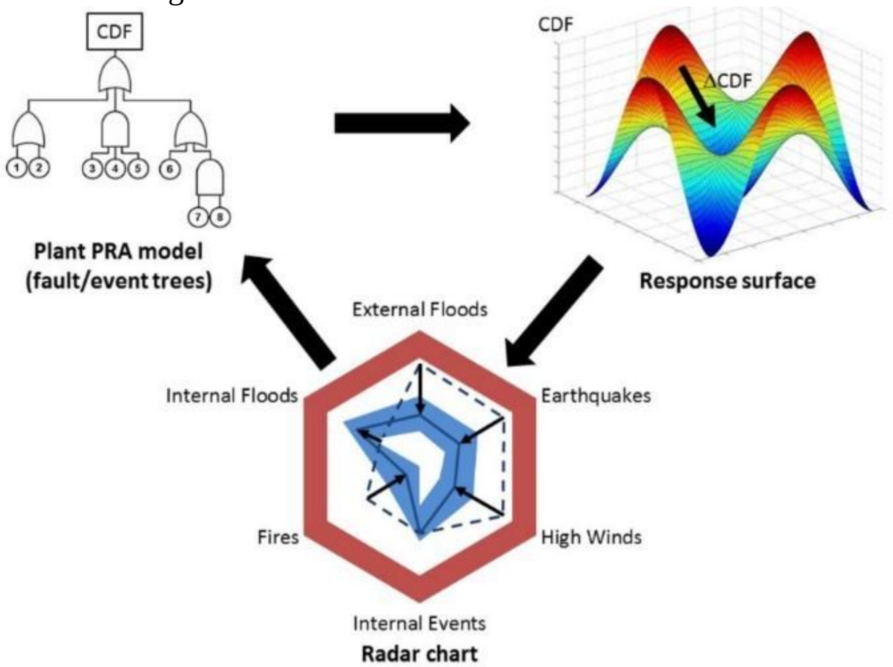

Figure 10 Iterative process to "aggregate"

The iterative process using together PSA, RSM and radar charts is proposed as a process to aggregate the results from different hazard groups. But actually, the use of RSM only does not really allow to identify all the significant contributors. It does not take into account the "hidden" uncertainties which lie, for example, in the assumptions we make when we lack of knowledge. And radar charts do not really allow the aggregation of the hazards from different hazard groups. They only make decision makers able to compare the relative result of each hazard group (associated with its uncertainty) to a given threshold. There is an implicit rule for decision making with this type of representation which may not be applicable in every application.

4. Goal of the research

Figure 11 gives a brief illustration of the main goal and its benefits, where we are seeking to introduce a way and creating a comprehensive model that represents the communication gate in the figure below, which is able to assess the level of heterogeneity of risk assessment for each hazard group (given the lego consisting parts of the final metrics, the uncertainty, the available knowledge, the model and methods used, etc.), which should give a sense of realism of the risk prediction, and to be easily understandable and accessible by the decision makers, in order to help improving the decision making process. In other language we should create a common frame for both the analysts and the decision maker that covers the risk and guarantees the proper decision, which should lead to safer practice of the technology.

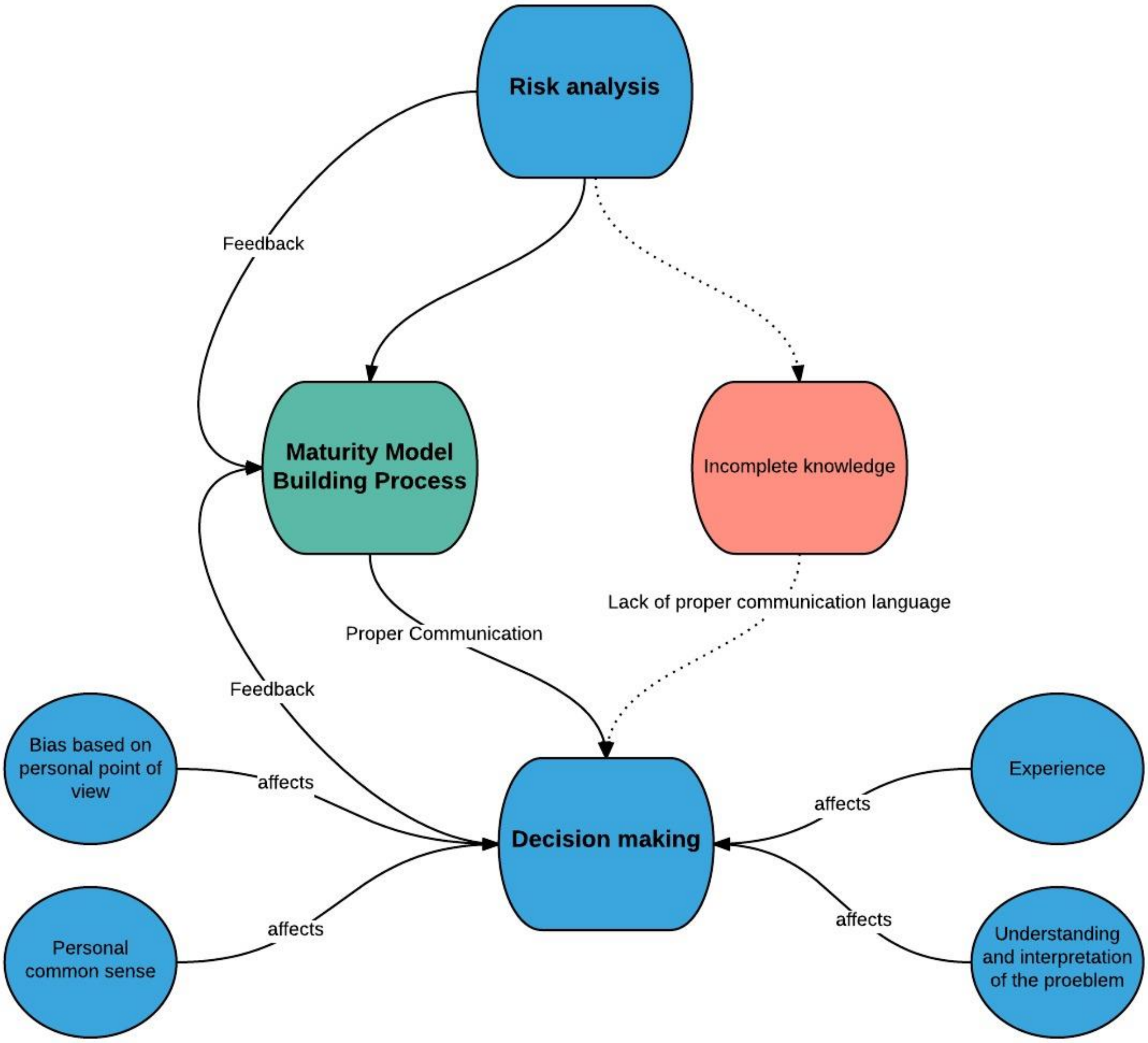

Figure 11 Risk informed decision making process with and without level of maturity model.

5. First developments

As we explained before, our goal is to create a model that should cover all the drawbacks of the usual techniques that is used for risk aggregation in risk-informed decision making, in which it is able to unify all the bases of different hazard group, and overcome the heterogeneity whenever we proceed to make the aggregation, as well as building a common base which is understood by both the analysts and the decision maker. Actually we thought about introducing what we call the level of maturity, which can be defined as the level of analysis, knowledge, capability and trustworthiness of the model taking into account different elements that assess the reliability of the model and its capability of prediction; in other words the assessment of how much we can trust our model and analysis.
Actually a similar idea was introduced and explained in a paper regarding the predictive capability maturity model for computational modeling and simulation [10], which aims to cover the following:

- The model should be able to assess the usefulness of the analysis in order to better inform and improve decision making
- It should help in increasing the adequacy of prediction to meet accuracy requirement for system response of interest.

In this model they have addressed six different criteria (contributors) to the level of maturity, where from their point of view they believe that those different criteria affects the level of maturity of the model, as illustrated in ***Table 1***.

Table 1 Example of PCMM Table Assessment and Project Maturity Requirements [10].

MATURITY / ELEMENT	Maturity Level 0	Maturity Level 1	Maturity Level 2	Maturity Level 3
Representation and Geometric Fidelity		Assessed	Required	
Physics and Material Model Fidelity			Assessed Required	
Code Verification		Assessed Required		
Solution Verification	Assessed		Required	
Model Validation		Assessed	Required	
Uncertainty Quantification and Sensitivity Analysis	Assessed			Required

They gave different colors for each element to address if they meet the requirement of maturity or not, you can notice that the scores of the different criteria are [1,2,1,0,1,1] while what is required is [2,2,1,2,2,3]. The next step which might be more challenging is the aggregation, especially that it is known that collecting information for decision maker can be really challenging and difficult. For this model they have suggested a very simple way of aggregation in order to help maintaining the key information in the different individual scores as the following ([10]):

$$\widehat{\mathrm{PCMM}} = \left[\min_{i=1,2,\dots n}\mathrm{PCMM}_i, \frac{1}{n}\sum_{i=1}^{n}\mathrm{PCMM}_i, \max_{i=1,2,\dots n}\mathrm{PCMM}_i\right], \qquad (1)$$

In this way a set of three scores are computed and presented to the user (decision maker). It represents the minimum, the average, and the maximum of the aggregation of the elements.
An example that was presented by the same paper:

$$\mathrm{PCMM}_{subsys1} = \begin{bmatrix} 1 \\ 1.5 \\ 1 \\ 0 \\ 0.5 \\ 1 \end{bmatrix}, \mathrm{PCMM}_{subsys2} = \begin{bmatrix} 1.5 \\ 1 \\ 0 \\ 0.5 \\ 1.5 \\ 0 \end{bmatrix}, \mathrm{PCMM}_{subsys3} = \begin{bmatrix} 2 \\ 1.5 \\ 0.5 \\ 1 \\ 1.5 \\ 1 \end{bmatrix}, \mathrm{PCMM}_{subsys4} = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0.5 \\ 1.5 \\ 1.5 \end{bmatrix} \qquad (2)$$

And using equations 1 and 2 they computed the PCMM aggregate triple:

$$\widehat{\mathrm{PCMM}} = [0.0, 1.1, 2.0] \qquad (3)$$

In a similar way we are thinking about looking for the different criteria that we believe they can affect the level of maturity. The idea comprises of finding these criteria and giving a score for each depending on previous definition and classification, and then aggregation of the scores according to a mathematical model that we choose similarly to what was shown before in the PCMM model, to come out with a final result that represents a score of the level of maturity. This level of maturity must help the decision maker or the analyst to understand what should be the next step, whether through doing further analysis if needed, or through taking the decision relying on these data.
The maturity model can be aggregated in different methods, one of them is the suggested previous method to give triple results. Another method suggests the direct aggregation of each score from each criteria of each hazard group, similarly to what mentioned before but giving an average or a final summation of scores, which represents a level of satisfactory in terms of level of maturity. Each element (criteria) is to be given a specific score for each system in the hazard group, which results in a total sub-maturity model that can be aggregated of different hazard groups.

$$m_k = \sum_{j=1}^{N}\sum_{i=1}^{n}(\mathrm{MCr})_{i,j} \qquad (4)$$

Or using the average aggregation method:

$$(m_k)_{Avg} = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{n}(\mathrm{MCr})_{i,j} \qquad (5)$$

Where m_k: is the level of maturity of each hazard group, $(\mathrm{MCr})_{i,j}$ represent different criteria i for different systems j within each hazard group k. Furthermore the total level of maturity of the total analysis can be obtained by summing the different sub-maturity of each hazard group as illustrated in the following.

$$M = \sum_{k=1}^{R} m_k \qquad (6)$$

Or using the average aggregation method:

$$(M)_{Avg} = \frac{1}{R}\sum_{k=1}^{R}(m_k)_{Avg} \qquad (7)$$

The problem with this type of aggregation that it might be misleading for the decision maker in some cases if it is needed to know the causes of low level of maturity. A better way is to represent this kind of aggregation in matrices, and representing a final matrix of maturity that represents the scores of each criteria as illustrated in equation 6, which could be better informing for the decision maker, or it could be achieved by making the aggregation while keeping a symbol referring to each criteria as illustrated in equations 9 and 10.

$$\mathrm{M} = \sum_{k=1}^{R} \begin{bmatrix} \mathrm{MCr}_1 \\ \mathrm{MCr}_2 \\ \cdot \\ \cdot \\ \vdots \\ \mathrm{MCr}_n \end{bmatrix}_k = \left(\mathrm{m}_{\mathrm{hazgrp2}} = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 5 \\ 1 \\ 2 \end{bmatrix} \right) + \left(\mathrm{m}_{\mathrm{hazgrp2}} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \right) +. \dots \dots, \tag{8}$$

$$\mathrm{m}_k = \sum_{j=1}^{N} \sum_{i=1}^{n} (\mathrm{sc})_{i,j} (\mathrm{MCr})_{i,j} \tag{9}$$

And in vectors:

$$\mathrm{m}_{\mathrm{internal-fire}} = (\mathrm{sc})_1 (\mathrm{MCr})_{1,1} + (\mathrm{sc})_2 (\mathrm{MCr})_{1,2} + (\mathrm{sc})_3 (\mathrm{MCr})_{1,3} + \cdots \tag{10}$$

Or using the average aggregation method:

$$(\mathrm{m}_k)_{\mathrm{Avg}} = \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{n} (\mathrm{sc})_{i,j} (\mathrm{MCr})_{i,j} \tag{11}$$

Where $(\mathrm{sc})_{i,j}$ represents the scores in numbers and $(\mathrm{MCr})_{i,j}$ represents the criteria in symbols which helps the analyst in reviewing the analysis easily, as well as giving the decision maker the indication of the real meaning of the final score.

This method is based on similar ideas of MCDM (multiple criteria decision making) methodologies, where in MCDM methodology you need to measure and rate different options regarding a number of different criteria in order to compare them. Actually this method can be really very helpful in decision making process, where it can be used to rank different decision regarding the different criteria that have been set by both the decision maker and the analyst. This leads us to think about some tools such as **ELECTRE TRI.** ELECTRE TRI approach is concerned by multiple criteria decision making problems that are designed to address different problem sorting and segmentation, by assigning categories to pre-defined categories (criteria) [11]. Actually our future goal is to try to connect and integrate both the Maturity model approach and ELECTRE TRI approach into a comprehensive new methodology in order to enhance the decision making process.

So now and in order to judge the trustworthiness and the sensitivity of our analysis and as illustrated before, we should take into account different perspectives and approaches (criteria and element) that we believe it influences the risks aggregation's result a great deal. Some of these criteria that affect the trustworthiness and what we call the maturity level of the model are: the degree of our knowledge about this specific problem and model, the degree to which the models are being extrapolated from the real case to the condition of the application of interest; approximations made by the model and on the model, degree of precision and accuracy of our analysis, the degree to which we perform the analysis, and many other factors and criteria that will be illustrated in the following section.

Terje Aven pointed to the importance of some criteria and their link with risk in his article "conceptual framework for linking risk and the elements of the data–information–knowledge–wisdom (DIKW) hierarchy" [12], where he discussed how the risk is (which calculated only by multiplying the consequences by the frequency of occurrence) and those perspectives are linked to the knowledge and its different features. Other points might be of interest in evaluating the risk analysis models and their capability of predication through different metrics for evaluation of prognostic performance [13]. Based on different papers, we listed the model's elements that we believe it influences a lot the result as the following:

1. Certainty: it refers to the knowledge perfection, where through this criteria we should able to give a score regarding the uncertainty analysis that might be done quantitatively at the beginning and then qualitatively by giving it a rank or a score.
2. Level of importance of each consisting element of the hazard group: the importance analysis of the system's components, such as the importance analysis that is related reliability and availability requirements; ex: Birnbaum importance measures, and Fussell-Vesely importance measures [14].
3. Level of conservatism in the analysis: this criterion is a measure of the tendency of caution during the risk analysis that arises from concerns regarding the lack of data and knowledge about the nature and magnitude of the hazard.
4. The sensitivity of the analysis: sensitivity analysis is generally used to determine how a dependent variable can be changed and affected by the change of the input independent variable. This criterion is a measure of how the uncertainty in the output of a model can be affected or propagated regarding different sources of uncertainty in the input [15].This analysis can be done using response surface and cliff edges analysis [5].
5. Dependency: it is a rank of the dependency of one criterion on the other.
6. The type (criteria) and the level of trustworthiness in the analysis tool: it gives a score regarding the suitability of a specific model or risk analysis tool in a specific problem, and the previous feedback of the trustworthiness of this tool.
7. The state of knowledge regarding each specific hazard group, (can refer to the previous experience and our knowledge regarding a specific subject)
8. The level of analysis: ranking of the level of the details that are taken into account during the analysis.
9. Other algorithmic prognostics prediction assessment features (Accuracy and Precision): it measures and ranks the quality of data, and the characteristic of features.

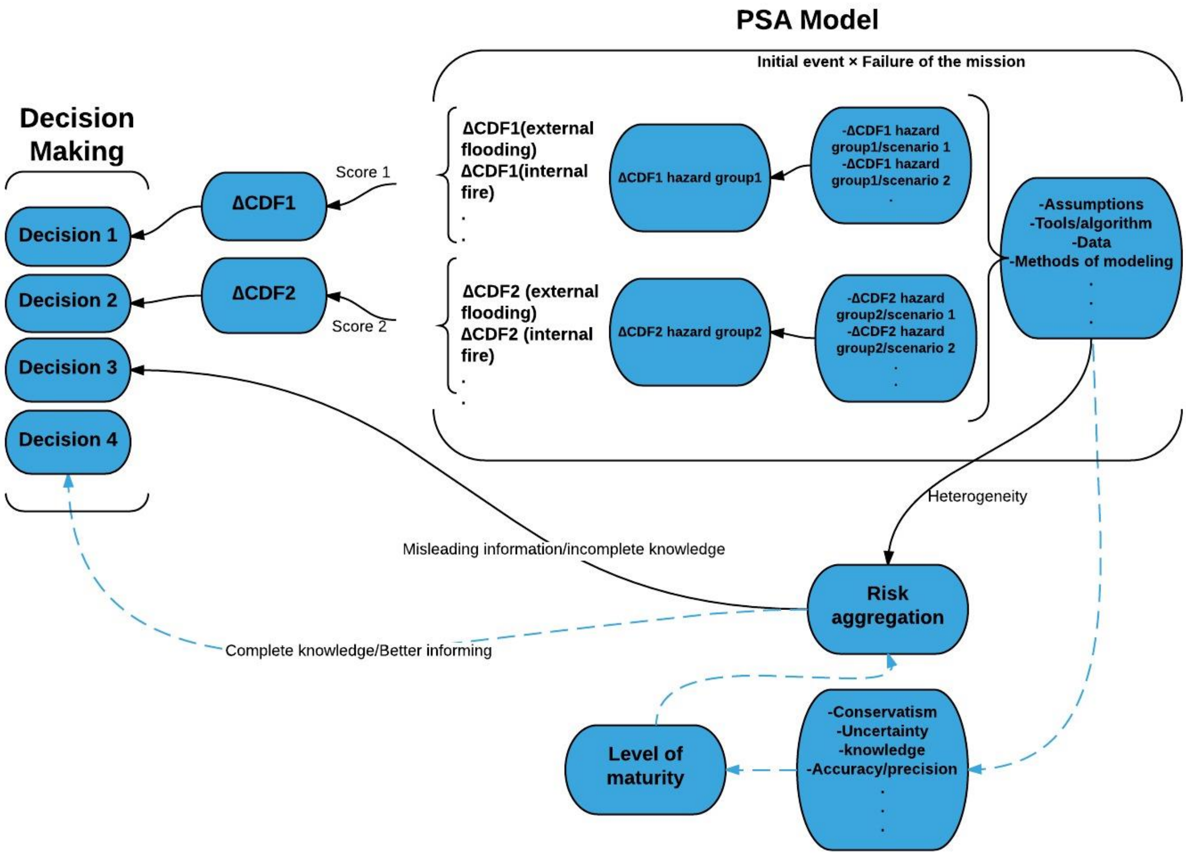

Figure 12 Risk analysis progression and steps with risk informed decision making with and without maturity model.

Actually ***Figure 12*** sums up the total work that has been done so far, where it is seen that usually in order to make a decision or to assess the benefits of a specific modification of the plant a final change on the metric is studied (such as ΔCDFmetric). The total metric or the total change on the metric is calculated with the assumption of different scenarios which leads the decision maker to make a specific decision regarding the final metric score. Actually the problem becomes more complicated and challenging regarding the risk informed decision making when the analyst needs to aggregate the total risk from different hazard groups, which is usually the case.

As the risk aggregation is needed to compare the final metric to a specific threshold value, the risk analysts usually do that using a simple summation which in the other hand presents difficulties in terms of interpretation of the results, given the current limitations of PSA including hazards. The figure above illustrates that using the level of maturity model the analysis should be better informing for the decision maker, as well as creating a proper communication gate between the decision makers and the analysts.

6. Criteria (elements) Definition and Scaling

Definition and scaling of the model's criteria, as well as establishing a guide for the analysts and the decision makers are essential to allows them to understand what are those criteria, how to give the scores for the different hazard groups, what is needed during the model building, and even during decision making process. Such a mission requires giving first definition and scaling with the help of the experts, and to benefit from their experience and common of sense regarding the different important details of the assessment that was developed during their work practice. Obtaining complete, accurate, and comprehensive definitions requires as well giving feedback during the primary test of the model on some examples, regarding the problem and the challenges that faces the analysts while giving the scores from the point of view of the precision of the definition.

As a first step the following points should be considered before giving the definitions:

1. The score is given quantitatively or qualitatively.
2. Scores are giving relative to each other.

It is suggested to give different relative scores for each one of these elements (ex; scores 1-5 for the uncertainty, depending on the level of the uncertainty), by giving a relative definition and description for each level and score

6.1 Rough classification of the levels of the model's elements

At the current moment we are going to give a brief rough classifications of each element of the maturity model, taking into account that these classification might vary from problem to another. Actually a good classification of each of these elements needs a deep study of these elements generally and the problem itself. Later on we will try to choose the best classification of the scores of each of these elements, making sure that no important information would be lost between these giving scores.

So as a starting point we choose to focus on 3 different elements (criteria) which are the knowledge, the level of conservatism, and the uncertainty. As a first step working on these 3 different elements, and relying on previous scaling and definition of these criteria we tried to give a clear definition and description for each element as it is illustrated in the following section, in order to help the analyst to in evaluating each level.

a. **The uncertainty** is defined by US EPA [16] as "the imperfect knowledge of the true value of a particular variable, or its real variability in an individual or a group". The uncertainty can be classified into different levels relatively, depending on the degree of knowledge perfection. What distinguishes the uncertainty in particular that it can be analyzed using different models, and obtain a final mathematical results that is sensed more easily due to its explicit nature, which can make the mission even easier when it is required to give a relative qualitative score of the degree of uncertainty.

Actually Hauke Riesch discussed the uncertainty level in one of the books regarding risk theories [17], where he gave a different scales for the uncertainty through a general definition of these levels as the following|:

- Level 1 of uncertainty (level 5 of certainty): Uncertainty about the outcome, where the model is known, the parameters are known, and it predicts a certain outcome with a probability P. (The traditional mathematical and philosophical problems of probability theory are most concerned with level 1.
- Level 2 of uncertainty (level 4 of certainty): Uncertainty about the parameters, where the model is known, but its parameters are not known, once the parameters

fixed then model predicts an outcome with parameters with probability P. (lack of empirical information, (ex. large break in primary circuit which has never happened).

- Level 3 of uncertainty (level 3 of certainty): Uncertainty about the model, where there are several models to choose from, and we have an idea of how likely each competing model is to reflect reality. (which model is more suitable for our problem)

- Level 4 of uncertainty (level 2 of certainty): Uncertainty about acknowledged inadequacies and our implicitly made assumptions. Unmolded uncertainty, and their estimation of their uncertainty will be different according to background knowledge and assumptions.

- Level 5 of uncertainty (level 1 of certainty): Uncertainty about unknown inadequacies and it is corresponding to unforeseen events, non-modeled and non-modable risks. In other language it is the uncertainty when we do not even know what we don’t know (unknown unknowns).These limitations could arise because some aspects that we know of have been omitted, or because of extrapolations from data or limitations in the computations, or a host of other possible reasons. A slightly more formal way of responding to unforeseen events is the introduction of ‘‘fudge factors,’’

b. **Knowledge:** refers to the amount and the type of data, the suitability and the applicability of this data corresponding to the problem, the ability to apply and utilize these data, the sources and the origins of these data and its relation with the suitability and trustworthy, the type of models used to obtain or use these data, the cognition regarding the type of the technology and each specific system, the experience in this technology, the level of the experience of the analyst, the type of knowledge (explicit, implicit, tacit), the background knowledge.

1. The degree to which data/information and are available and reliable/relevant

2. The degree to which the assumptions are reasonable/realistic

3. The degree to which there is agreement among experts

4. The degree to which the phenomenon involved are understood and accurate models exist.

Actually as it is known that the knowledge itself has so many aspects to be considered when it is needed to give the score, it is better to give the score of the level of the knowledge relying on previous questions imposed. The imposed questions can have multiple answers that is defined previously and each one of these answers refers to a specific score. Table 2 gives an example on of some questions that can be imposed.

Table 2 Knowledge Criteria levels definition and description

Knowledge criteria	Level 1	Level 2	Level 3	Level 4	Level 5
Amount of data sources, references	No data	The data are so limited and can be extracted only from the same type of nuclear power plants	The data are available and can be extracted from any other nuclear power plant	The data are available in all kinds of power plants	The data are Available in abundance and can be extracted or gotten easily from so many sources and places world wide
Type of needed data	General and can be extracted from any place	General for all kind of power plants	Specific for nuclear power plant	Specific for types of reactors, (PWR, BWR, CANDUM….)	Specific to your plant
Source of data	Other sources that is not related directly to the technology	Other power plants and other	Other nuclear power plants	Other nuclear power plants of the same type and technology	Own plant
Applicability and suitability for the case study	Not really applicable and suitable for the NPP	Applicable and suitable for power plants	Applicable and suitable for NPP generally	The data are applicable for our NPP, but not perfectly applicable to our case	The data are perfectly suitable and applicable to our case
Quality of the data and the models	Assumption of data	Data are calculated using statistical models	Data are both assumed and calculated using computer physical and mathematical models	Data are gotten using computer mathematical and physical models	The data are measured precisely and accurately, and then modeled
Experience in the initiating event	No experience at all	Experienced such an event in other industries	Experienced such an event in nuclear power plants industry	Experienced this event in the same nuclear power plant	This event is quite common and we have a wide experience inn

Experience in the technology	The technology is unknown and applied for the first time in all industries	The technology is known only by nuclear industry	The technology is known worldwide but applied for the first time in in nuclear power plants industry	The technology is known worldwide and used in many nuclear power plants but applied for the first time in this power plant	The technology is very well known worldwide in all kind of nuclear power plants
Level of the experience of the analyst As we believe that the opinions are influenced by the knowledge and experience	He is a risk analyzer but has no experience at all in this specific domain of nuclear power plants	It is his/her study domain and he/she has training courses before starting the real analysis	It is his specialty and he practiced through training courses regarding the same type of nuclear power plants, and had training courses directly before the analysis taking into account all the newly needed updates	He is specialized in the very specific domain that he is going to work on (safeguard system, auxiliary system) and he practiced this kind of analysis before	Expert in this specific domain and had done this work before
Background knowledge and its life time and trustworthiness of the source	We have information that is not proven	The data are got from other international companies and it is proven by the company itself	The data are gotten from nuclear power plants specialized companies, it is a proven by match with data from other companies	The data are gotten from your company and it is proven by match with data from other international companies	The data are gotten from international nuclear agency such as IAEA and NRC and it is proven and approved by so many international companies
The match between the reality and the (gap between) how realistic the analysis are	The analysis are not realistic at all	There is only few realistic points in the analysis	The analysis are partially realistic	The analysis are realistic but do not match all real cases that took place before	The analysis are so realistic and matches a lot most of the real cases that took place

The conservatism arises from the fact that it is “Better safe than sorry” Samuel Lover’s Rory O’More (1837). Some has referred to the conservatism in risk management as the desire for caution that rises from different consideration and reasons such as the concerns regarding the lack of knowledge about the nature and magnitude of the hazard [18]. Others refer to the conservatism as a preference to make a mistake in the side of overestimating the risk rather than understanding it under conditions of uncertainty that might be underestimating of it. This tendency is manifested when risk estimation exceeds the mean value of risk probability distribution (that is neither underestimating, not overestimating the risk); where conservatism might involve selecting a risk estimate at for example 95th percentile, or 99th percentile, meaning there is a 99% probability that the risk is over estimated and 1% is underestimated [19].

Actually at the beginning we thought about classifying the levels of conservatism as the following:

1. Extremely conservative: the level of knowledge is very low and we have a high level of uncertainty.

2. Conservative : we need to be conservative due to the lack of knowledge and the high level of uncertainty

3. Modest conservatism : we have a very good level of knowledge, with quite low uncertainty, quite good reliability

4. Very modest conservatism: when we have a high level of knowledge, the uncertainty is quite low, the reliability is very good, and the importance is low.

5. No conservatism: the knowledge is absolute and there is no uncertainty, the reliability is high, and the importance (the criticality of the rule of the part or the detriment in case of the fail) is low, so we do not need to be conservative

According to the previous we can see the dependency of this element (conservatism) with other criteria such as the uncertainty and importance.

Actually we tried to enhance the definition of the levels as we have found that it is not fitting with the different hazard groups.

1. I’m sure to be conservative in a way that covers even the unknown unknowns
2. I’m conservative and I know it
3. I think I’m conservative but I’m not sure
4. I think I’m not conservative
5. I’m sure that I’m not conservative.

7. Final Mathematical Model of the Level of Maturity

As we discussed before the final model of maturity is a function of several criteria that affect the analysis and the degree of trustworthiness of the model. Integrating the different maturity criteria into the maturity model results in the following:

Maturity level model = {uncertainty, importance, conservatism, sensitivity, dependency, trustworthiness of the analysis tool, state of knowledge, level of analysis, accuracy, precision}

In symbols:

$$m_j = \{(sc)_1 u_j, (sc)_2 i_j, (sc)_3 c_j, (sc)_4 s_j, (sc)_5 d_j, (sc)_6 t_j, (sc)_7 k_j, (sc)_8 a_j, (sc)_9 ac_j, (sc)_{10} p_j\} \quad (12)$$

$$M = \{(SC)_1 U,\ (SC)_2 I,\ (SC)_3 C,\ (SC)_4 S,\ (SC)_5 D,\ (SC)_6 T,\ (SC)_7 K,\ (SC)_8 A,\ (SC)_9 Ac,\ (SC)_{10} P\} \quad (13)$$

ex; $m_{\text{internal fire}} = \{3u,\ 3i,\ 4c, \ldots\ldots\ldots..\}$ means that the internal fire hazard group scores uncertainty of level 3, importance of level 3, and conservatism of level, which can be also represented in matrices as previously. Actually we kept the symbol of each element of the maturity model, in order to keep a sense of the meaning of these number before the final step of scalar aggregation, which would be better informing for the decision maker.

$$M = m_{\text{internal fire}} + m_{\text{internal flooding}} + m_{\text{external flooding}} + \cdots. \quad (14)$$

$$M = \begin{bmatrix} C \\ Cv \\ K \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix} + \cdots.. \quad (15)$$

And supposing that there are only 3 hazard groups $M = \begin{bmatrix} 10 \\ 7 \\ 8 \end{bmatrix}$ in the direct summation method, and $M = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$ in average aggregation method.

Another suggestion is to introduce the risk as well in the same model, which might at the hazard group level or at the sub-hazard group (consisting elements of each hazard group)

Hazard group level: $m_j = \text{Risk} \times$
$$\{(sc)_1 u_j, (sc)_2 i_j, (sc)_3 c_j, (sc)_4 s_j, (sc)_5 d_j, (sc)_6 t_j, (sc)_7 k_j, (sc)_/ a_j, (sc)_9 ac_j, (sc)_{10} p,\} \quad (16)$$

Suggestion:

- In this type of model it is suggested to keep the symbols to give a clue about the scores meaning.

8. Conclusion

As discussed through this report, the risk aggregation is a very important process in risk informed decision making, as it is needed to compare a final metric of risk; (core damage frequency, early large release) to a reference or a threshold value. In fact risk aggregation using classic methods might be misleading for the decision maker, as it is usually achieved using a simple summation of the risk from different hazard groups which is not mathematically consistent nor physically meaningful due to the heterogeneity of different hazard groups.

Our idea was to create what we called the level of maturity model, which should be able to overcome the problem of hazard group heterogeneity. The model consists of different criteria (elements) that are believed to affect the level of maturity of each hazard group such as: the level of uncertainty, the level of conservatism, the level of knowledge, and etc.

As a start the model was applied on a basic example (look at table A.1 in appendix A) using only 3 criteria, and as it is believed that that the score (rank) won't be precise unless so many aspects of each single criteria are taken into account, we have started by dividing the level of knowledge into different other sub- criterion, and the final score of the level of the knowledge was given by averaging the different scores of each sub-criterion. Actually giving a precise definition and sub-criterion for criteria is what we are planning to do in the near future for the PhD purposes, especially that we noticed the need to give a more specific and precise definition or even to change the way of scaling during the application of the model on some hazard groups to make more compatible.

The application of this model on a basic example, has opened the eyes on other important and challenging point. In fact It is not really convincing to use all the criteria of the model in the same way, because it is believed that each criteria contributes in different way and percentage in the level of maturity, which is usually done in multiple criteria decision making MCDM process, as well as in ELECTRE TRI approach with the help of the management. In the other hand for maturity level model, the weighting of each criterion should be accomplished with the help of the experts (the analysts and the decision maker), using their experience and common sense to give an estimation of their contribution, where each criterion is then weighted, according to how important it is in the overall picture, relative to all the rest.

The model of maturity is not complete yet, and it still needs a lot of developments regarding the mathematical model, the different perspectives of the definition of each criterion and their contribution, as well as trying to integrate known multi criteria decision making approaches (ex. ELECTRE TRI approach in order to use it in ranking the different possible combination of scores) with our model to enhance it, and finally to make a guideline for both the analysts and the decision maker to help them understand and apply better the model, which are believed to be our goals for the PhD thesis.

References

[1], N. (1995). Use of probabilistc Risk Assessment Methods in Nucear Regulatory Activities; Final Policy statment.

[2], N. S. (n.d.). FIRE PRA MATURITY AND REALISM: A DISCUSSION AND SUGGESTIONS FOR IMPROVEMENT.

[3], M. D.-G. (n.d.). NUREG-1855 (Guidance on the Treatment of Uncertainties Associated with PRAs in Risk-Informed Decision Making). NRC.

[4], M. (2012). EPRI Technical Update, Practical Guidance on the Use of Probabilistic Risk Assessment in Risk-Informed Applications with a Focus on the treatment of Uncertanity. EPRA.

[5], N. Z.-P. (n.d.). Addressing multi-hazards risk aggregation for nuclear power plants through response surface and risk visualization.

[6], N.-2.-5. (2010). NASA Risk-Informed Decision MakingHnadbook.

[7], T. A. (2003). On the use of risk and decision analysis to support decision making.

[8], O. A. (2015). On how to Understand and acknowledge risk.

[9], EPRI Technical Update. (2014). An Approach to Risk Aggregation for RisK-Informed Decision-Making.

[10], W. L. (2007). Predictive Capability Maturity Model for Computational Modeling and Simulation.

[11], V. M. (2000). A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support.

[12], T. A. (2013). A conceptual framework for linking risk and the elements of the data–information–knowledge–wisdom (DIKW) hierarchy.

[13], NASA. (2010). Metrics for Offline Evaluation of Prognostic Performance.

[14], E. Z. (2009). Computational Methods for Reliability and Risk Analysis vol 14.

[15], S. A. (2008). Global Sensitivity Analysis.

[16], C. E. (1999). Draft Hazardous Waste Combustion Protocols - Risk Characterization | US EPA ARCHIVE DOCUMENT.

[17], S. R. (2013). Essentials of Risk Theory.

[18], W. K. (1997). Conservative versus Mean Risk Assessments: Implications for Superfund Policies.

[19], R. M. (1996). comment on: Does Risk Aversion Make a Case for Conservatism? RISK: Health, Safety and Env't.

APPENDICES

APPENDIX A

Table A.1 Example of the model of Maturity application

Certainty	Conservatism	Knowledge	Example	comments
1	1	1	It does not exist in our model (not realistic)	All combination of certainty 1 do not exist in our model, giving the definition of level one of certainty
2	3	1	Large LOCA frequency	We gave 2 for the certainty because we do not have any model and data. I'm not sure of being conservative or not, and I have no background knowledge regarding large LOCA.
4	2	3	Earthquake frequency for a giving level	We know the model, but we are uncertain about the input parameters, and we have data but it is for a wider region in general not for a giving site, according to the expert they are conservative, we have a quite good level of knowledge as we have a background data and we have models and experimental data.
3	2	3	External flooding frequency	The uncertainty is 3, as we are not the certain about the input parameters, neither the models. The level of conservatism is 2 because we are conservative and we know it, because we take into account past events that could not happen, and using our model and comparing them to more realistic models (Shadex), we found out that model is conservative (ISERE). Knowledge: we put 3 which the average of 4 different knowledge aspects (the amount of data :2, source of data :4, applicability of data: 4, and the quality of the model: 2)

4	2	3	Internal fire frequency	We have quite good amount of data : source of our data: 4, amount of data: 3, applicability: 4, quality of data: 2, the model: 4 so the knowledge is : 3,
4	3	4	Loss of auxiliary feed water system frequency , loss components cooling system	Source of the data: 4, amount of the data: 2, type of data: 4, quality of data: 4, Models: 4, so knowledge: 4. Conservatism: we do not take into account passive redundancy, we do not take into account repairs, we use minimal cut set.
4	3	3	Small LOCA frequency	Knowledge: source of knowledge: 3, amount of data: 3, type of data: 3, model: 3. I think I'm conservative because I get the data from other nuclear power plants which means that we are taking into account too much events, but we are not sure because we do not take into account other phenomenon such as aging.
3	3	2	Internal flooding frequency	Knowledge: source of data: 3, amount of data: 1, model: 1. Conservatism: 3 we think we are (we use a lot of operating data, but we are not sure if they are applicable)