Development of good practice guidance for quantification of thermal-hydraulic code model input uncertainty

J. Baccou, J. Zhang, P. Fillion, G. Damblin, A. Petruzzi, R. Mendizábal, F. Reventós, T. Skorek, M. Couplet, B. Iooss, et al.

Development of good practice guidance for quantification of thermal-hydraulic code model input uncertainty

Jean Baccou[1]*, Jinzhao Zhang[2], Philippe Fillion[3], Guillaume Damblin[3], Alessandro Petruzzi[4], Rafael Mendizábal[5], Francesc Reventós[6], Tomasz Skorek[7], Mathieu Couplet[8], Bertrand Iooss[8], Deog-Yeon Oh[9], Takeshi Takeda[10]

[1] Institut de Radioprotection et de Sûreté nucléaire (IRSN), PSN-RES/SEMIA, Centre de Cadarache, 13115 St Paul-Lez- Durance, France; jean.baccou@irsn.fr

[2] Tractebel (ENGIE), Boulevard Simon Bolivar 34-36, 1000 Brussels, Belgium; jinzhao.zhang@tractebel.engie.com

[3] CEA, Université Paris Saclay, DEN/DM2S/STMF/LMES, F-91191 Gif-sur-Yvette, France; philippe.fillion@cea.fr; guillaume.damblin@cea.fr

[4] N.IN.E. - Nuclear and INdustrial Engineering S.r.l., Via della Chiesa XXXII, 759 - 55100 Lucca, Italy; a.petruzzi@nineeng.com

[5] Consejo de Seguridad Nuclear (CSN), Pedro Justo Dorado Dellmans, 11, 28040 Madrid, Spain; rmsanz@csn.es

[6] Universitat Politècnica de Catalunya (UPC), Avda. Diagonal 647, 08028 Barcelona, Spain; francesc.reventos@upc.edu

[7] Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) GmbH, Forschungszentrum , 85748 Garching, Germany; Tomasz.Skorek@grs.de

[8] EDF R&D, 6 Quai Watier, 78401, Chatou, France; mathieu.couplet@edf.fr; bertrand.iooss@edf.fr

[9] Korea Institute of Nuclear Safety, 62 Gwahak-ro, Yusong-gu, Daejeon 34142, Korea; k392ody@kins.re.kr

[10] Nuclear Regulation Authority 1-9-9, Roppongi, Minato-ku, 106-8450 Tokyo, Japan; takeda.takeshi4695@gmail.com

* Correspondence: jean.baccou@irsn.fr

Received: date; Accepted: date; Published: date

Abstract:

Taking into account uncertainties is a key issue in nuclear power plant safety analysis using best estimate plus uncertainty methodologies. It involves two main types of treatment depending on the variables of interest: input parameters or system response quantity. The OECD/NEA PREMIUM project devoted to the first type of variables has shown that inverse methods for input uncertainty quantification can exhibit strong user-effect. One of the main reasons was the lack of a clear guidance to perform a reliable analysis. This work is precisely devoted to the development of a first good practice guidance document for quantification of thermal-hydraulic code model input uncertainty. The developments have been done in the framework of the OECD/NEA SAPIUM project (January 2017-September 2019). This paper provides a summary of the main project outcome. Recommendations and open issues for future developments are also given.

Keywords: good practice guidance**,** model input uncertainty quantification, system approach, thermal hydraulic code, validation, inverse quantification of uncertainty.

1. Introduction

Assessment of uncertainties associated with Best-Estimate (BE) calculations has become a key issue in the so-called Best-Estimate Plus Uncertainty (BEPU) nuclear safety analyses. The question of input uncertainty forward propagation (i.e. the uncertainties on input parameters previously determined are propagated through the simulation model (SM) to estimate uncertainties on System Response Quantities (SRQs)) has been already addressed by several OECD/NEA projects (OECD/NEA, 1998; OECD/NEA, 2011). While it appeared that associated output uncertainty quantification methods have now become mature for industrial applications, it was also emphasized that a special attention should be devoted to the input uncertainty quantification (IUQ) on the physical models.

This paper is related to this last topic and more precisely to model input uncertainty quantification based on inverse propagation (IP) of the information associated to the discrepancy between simulation results and experimental data, using verified and validated simulation models. It concerns the development of a new systematic step by step approach for transparent and rigorous model IUQ. The construction has been performed in the framework of the SAPIUM project (Baccou et al., 2018) exploiting the lesson learned from the PREMIUM activity (Mendizábal et al., 2017) (a previous OECD/NEA benchmark devoted to IUQ) and a first investigation (Baccou et al., 2017) that led to the identification of five key generic elements that should structure the systematic approach.

This type of development has been proposed in order to minimize (or at least allow to understand) the user effect identified in previous projects. Moreover, by clearly compiling the different approaches as well as the methodological tools to handle each step, the SAPIUM framework offers a shared understanding about "appropriate" practices for model input uncertainty quantification in order to improve the reliability of the BEPU analysis and to progress on the validity of extrapolation of its results to the NPP case. Therefore, the main outcome of this work is a first good practice guidance document that can be exploited for safety study to increase the agreement among experts on recommended practices as well as on remaining open issues for further developments. The field of applications mainly concerns thermal-hydraulic model but the generic procedure as well as some of the described tools can be used for other types of models.

The paper is organized as follows. Section 2 is devoted to a summary of the good practice guidance document. Recommendations are then formulated in Section 3 to ensure a reliable model IUQ. Finally, we identify in Section 4 remaining key open issues that can be tackled in future developments.

2. Description of the systematic approach

2.1. Elements and major steps of the systematic approach

The good practice guidance is structured following 5 elements (Figure 1).

Figure 1. Elements of the SAPIUM approach.

In this systematic approach, Element 1 (specification) is common to any kinds of nuclear safety analysis studies. Elements 2 (experimental database) and 3 (simulation model) provide the information for the model input uncertainty quantification and validation. Interactions between them (depicted by the two-sided arrow on Figure 1) are required for their construction. Element 1 might also require using simulation model and sensitivity analysis tools (to confirm a PIRT (Phenomena Identification and Ranking Table) for example) addressed in Element 3. Since model IUQ methods are based on the comparison between simulation model results and experimental values, Elements 2 and 3 are crucial for their development. They will control the reliability of the final model input uncertainties and the capability of the method to extrapolate the results to real situations. Element 4 consists in inferring, from the comparison between simulation model calculations and experimental values, the information related to model input uncertainties. Finally, the validation performed in Element 5 is based on the propagation of all input uncertainties (integrating the quantified model input uncertainties obtained in Element 4) through the simulation model. It also exploits the experimental database identified in Element 2. This last element is interacting with Element 4 (depicted by the two-sided arrow on Figure 1) and might be used in a iteration process with Elements 2 to 4 (as indicated by the blue arrow) in case of non-acceptable validation results.

For efficient applications, each element is further split in different steps that are summarized in the following sections.

2.2. Element 1 - specification of the problem and requirements

As part of the BEPU methodology, the quantification of model input uncertainty should start with a clear and concise specification of the problem. It consists in the identification and definition of the SRQs and important (or key) physical phenomena for the intended applications, which are the first steps of the BEPU methodology for transient and accident analyses (IAEA, 2009; USNRC, 1989; USNRC, 2005), such as those included in the final safety analysis report (FSAR) (USNRC, 1978). This element provides recommendations and good practices for specifying the model input uncertainty quantification and validation problem in the following three steps (Figure 2).

Figure 2. Steps of Element 1.

2.2.1. Step 1 - Specification of the model IUQ purpose

In this Step, the type of nuclear power plant (NPP) and the transient or accident scenario of interest are first specified. Specification of the model IUQ objective is important because any given NPP type or transient scenario may be analyzed for different reasons. The specification influences the whole process of simulation model development, assessment, and specific analysis for each specific transient scenario.

The application domain or the applicability of the model IUQ is transient scenario-dependent because the dominant safety parameters and acceptance criteria differ from one scenario to another. Therefore, the transient scenario determines also the key phenomena and input uncertainties that must be quantified and validated. However, a complete scenario definition is NPP type (e.g., PWRs, BWRs, or CANDU) specific or sometimes even plant specific, because the dominant physical phenomena and their interactions differ in various reactor design or specific plant configuration.

It is thus recommended to start the specifications according to the transient classification as documented in the FSAR of the targeted NPP (USNRC, 1978), or in case of new plant design, to specify the application domain for the newly identified transient scenarios. In order to reduce the model IUQ efforts for a simulation model for a specific application, it is recommended to group different transient scenarios into a category of transients with common SRQs, and to make a generic model IUQ and validation for that class of transients.

2.2.2. Step 2 - Selection of system responses quantities

The selection of the SRQs should be made according to the objective of the model IUQ study. For model IUQ purpose, the chosen SRQs should be the parameters that are directly and accurately measured in the experiments. In most cases, this selection is straightforward, as they are directly related to the safety or design parameters in the acceptance criteria as specified in the applicable rules, guides, design codes or standards for the transient scenario of interest.

In nuclear power plant FSAR (USNRC, 1978), the acceptance criteria for the events of interest are defined in terms of quantitative fuel and reactor system design limits, such as reactor coolant system pressure or temperature limits, departure from nucleate boiling ratio limits, cladding or fuel temperature limits, etc. (ANSI, 1973, 1988). Thus, for FSAR accident analysis, SRQs are generally synonymous with criteria directly associated with the regulations, and their selection is usually a simple matter.

During simulation model development and assessment, a surrogate variable (i.e. a replacement of the SRQ by another measurable or predictable variable) may be of values in evaluating the importance of phenomena and processes. In such a case, justification for using a surrogate SRQ should be provided. In line with the surrogate SRQ, it is also important to consider other related performance measures in conjunction with the principle objectives. Because compensating errors in the simulation model can unintentionally lead to correct answers, additional performance measures serve as physical tracking points and additional proof of accuracy.

2.2.3. Step 3 - Identification of important phenomena (PIRT)

The involved physical phenomena for the concerned reactor components and systems and the transient or accident scenario of interest are then identified. Since many physical phenomena and input parameters may be involved in any thermal hydraulic analysis, and they are not modelled in a simulation model (or computer code) at the same level of fidelity, it may be impractical to quantify the uncertainty for each phenomenon and each input parameter. Therefore, it is essential to rank the importance of the involved physical phenomena or input parameters.

The behavior of a specific plant and scenario is not equally influenced by all the processes and phenomena that occur during a transient period. The most cost-effective but sufficient analysis reduces all potential phenomena to a manageable set by identifying and prioritizing phenomena according to their influence on the SRQs (Aksan et al., 2018). Each phase of the transient scenario and the system components are investigated separately. The processes and phenomena associated with each component are examined. The cause and the effect are differentiated. Once the processes and phenomena are identified, they are ranked according to their impact on the relevant SRQs, mostly based on the expert judgement. The main product of the process described above is the so-called Phenomena Identification and Ranking Table (PIRT) (Diamond, 2006; Wilson and Boyack, 1998). The formality and complexity of this process should be coherent with the complexity and importance of the scenario and component under consideration. In order to avoid the subjectivity of the expert judgement, some quantitative PIRT techniques (Luo et al., 2010; Martin, 2011; Yurko and Jacopo, 2012) could be used to confirm the importance of the identified phenomena, based on Global Sensitivity Analysis (GSA).

The development and assessment of a simulation model, including the model IUQ activities, should be based on a credible PIRT. The PIRT should be used to determine requirements for physical model or methodology development, scalability, validation, and sensitivity studies. In the end, the PIRT is used to guide any uncertainty analysis or to assess the overall adequacy of the evaluation model.

2.3. Element 2 - development and assessment of the experimental database

The objective of Element 2 is the construction of an adequate experimental database for the problem specified in Element 1. The construction of the experimental database relies on an efficient selection procedure of suitable experiments and associated tests from a large available database that can come from the test matrices used for thermal-hydraulic codes verification and validation. In order to answer these issues, the three following steps (4-6) are identified in the SAPIUM approach (Figure 3).

Figure 3. Steps of Element 2.

2.3.1. Step 4 - Establishment of a list of the available experiments and standardized description of each experiment

The experiments in nuclear thermal-hydraulics are usually split into several categories: basic tests, Separate-Effect Tests (SETs), Combined Effect Tests (CETs), and Integral-Effect Tests (IETs). SETs deal usually with one phenomenon whereas IETs are designed to investigate the overall system behaviors and the related phenomena and processes. Combined Effect Tests (CETs) are also identified with several interacted phenomena. These tests usually represent several components of a reactor but not the whole system, as the IETs. It may have a lower capability to give information on models than SETs and a lower capability to simulate all system effects but it may address coupled (combined) phenomena in a prototypical geometry (Mascari et al., 2015).

One possible strategy is to perform the quantification of the input uncertainty of physical models independently from the reactor transient scenario,. However, such strategy is practically impossible to address all possible situations with a unique set of model uncertainties without being systematically over-conservative. In practice, a scenario-dependent experimental database should be built according to the important phenomena from a PIRT, as suggested in Element 1. It has to be specified at the very beginning how detailed the subdivision of the experimental data base is necessary: parts of transient, different geometry or locations in the facility, presence of specific phenomena. It has to be checked if there are enough experimental data and if variations in the code of each uncertain parameter can be easily realized.

In order to help at selecting an experiment, a proposal of the standardized description of a test is given in the SAPIUM guide. It provides information on the type of the test: covered phenomena, experimental conditions, instrumentation, available measurements, and so on, which can be used as criteria for the selection. Experimental uncertainties, due for example to inherent limitations of the measurement technique or space and time resolution of the measurement, are one of the criteria to retain in a database. The experimental uncertainties of relevant measurements should be evaluated and provided.

2.3.2. Step 5 - Assessment of the adequacy of the database

The adequacy of the chosen experiment database must be assessed for the intended application. The adequacy of an experimental database includes two main properties that should be checked in the analysis. The first one is related to the ability of an experiment to provide relevant information for model input uncertainty quantification and validation. It is called representativeness. The second one concerns the ability of a set of experiments to fulfill the whole specifications of the problem under study e.g. cover the physical space of interest. It is referred as completeness.

The CSNI Code Validation Matrix (CCVM) of SETs and IETs for LOCA application and transients in LWRs gives a part of methodologies for the assessment of representativeness of an experiment and of completeness of an experimental database (Aksan et al., 1994; Annuziato et al., 1996).

The representativeness assessment first relies on the introduction of objective criteria to characterize the different tests of an experiment from the detailed description and with respect to the scenario. In the framework of model IUQ, priority for representativeness characterization should be given to: the phenomena-separable nature of the experiment, the quality of the experimental data, the agreement between experimental and case study conditions with a focus on the geometry and the Boundary and Initial Conditions, the capability of the experiment to address different simulation scales (CFD, component-scale, system scale). In practice, all the previous priorities are not always reachable. It is the case when there are only combined effect tests in the experimental database and no separated effect tests available. This problem was investigated within PREMIUM project and the answer comes from the development and application methodologies suitable for quantification of model input uncertainties on the basis of CETs (Skorek, 2017). Once criteria have been defined, representativeness evaluation can then be handled by exploiting Multi-Criteria Decision Making (MCDM) outranking approaches (see (Baccou et al., 2018) for an illustration) or Analytical Hierarchical Process (AHP) (Saaty, 1982) in order to objectively and automatically rank experiments.

Regarding the estimation of the completeness of the experimental database, criteria for assessing the maturity level of various issues within the V&V UQ procedure for nuclear licensing have been already proposed under the terminology Predictive Maturity Indexes (PMI) (Hemez et al., 2010). What emerges is that high level of maturity could contribute to prove the completeness of a database. However, since this SAPIUM element is restricted to the experimental database and not to the whole quantification process, the existing works on PMI need to be adapted. The ratio between the convex hull areas of the validation and application domains as well as the spatial distribution of the experiments within the validation domain can be relevant criteria to be taken into account in the construction of this index. Further discussions are provided in the SAPIUM report.

2.3.3. Step 6 - Selection of the experimental database for the model input uncertainty quantification and validation

The last step of Element 2 concerns the selection of the experimental database for the model input uncertainty quantification and the validation taking into account the adequacy assessment (Element 2, step 5). A classical strategy is application of SETs for quantification of model input uncertainties and validation of quantified model input uncertainties on the basis of IETs. The CETs were used in the past rather for validation than for quantification of model input uncertainties. This was mainly due to the fact that quantification of model input uncertainties on the basis of CETs, when several phenomena occurs simultaneously and several models input uncertainties have to be quantified in the same procedure, is a difficult task. However, with increasing accuracy of uncertainty analysis and improving methodologies, CETs are more and more used also for model input uncertainties quantification (Skorek, 2017). For some particular phenomena there are many tests (e.g. critical discharge), so the SETs can be applied for quantification as well as for validation. The usual situation is that for validation only IETs are applied. In the case of limited number of available experiments, the splitting of experimental data base is not useful. All the available experiments (including CETs and IETs) should be considered by quantification. The validation step can be performed by application of cross-validation procedure.

2.4. Element 3 - selection and assessment of the simulation model

The simulation model (SM) should be developed such as to adequately predict the SRQs for the transients or accidents of interest for the test facility (or the plant) from Element 2, and the key input parameters should be identified based on their importance to the modelling of the scenario and their impact on the SRQs for the simulation model calculation. The physical models included in the simulation model and their degree of fidelity in predicting physical phenomena must be consistent with the results of the PIRT process in Element 1. This Element 3 consists in the following 3 steps (Figure 4).

Figure 4. Steps of Element 3.

2.4.1. Step 7 - Selection of code based on capability assessment

The code for safety analysis should be chosen based on the investigated transient after a PIRT has identified the important phenomena for the given scenario and plant (see Element 1). The code applicability is determined based on the code user manuals,: the code's formulation, model, and correlations are ~~review~~ed to assess if the code has the model/correlations needed to simulate the important phenomena, and if the code has scale-up capability. It is recognized that while the formulation may be general, the correlations or constitutive relationships are empirical. The code's scalability will depend on these correlations and the underlying tests. If the tests scale the plant for the phenomena of interest, then correlation is applicable. However, if correlation was derived from tests that do not scale the plant, the code may not scale up the plant. Code validation with scaled tests or counterpart tests is another way of assessing code's ability for up-scaling and applicability to power plant (full) scale The analyses in following steps are addressed mainly to the system thermal-hydraulics codes such as RELAP5, TRACE, CATHARE2 (Petruzzi and D'Auria, 2008).

2.4.2. Step 8 - Development and assessment of applicability of the simulation model

2.4.2.1. Development of simulation model for all the tests of the experimental database

While the computer code and the uncertainty analysis methodology are the key elements of a BEPU approach, the nodalization strategy and model options used to simulate the test facility or plant, i.e., the simulation model, is also very important, as they connect the code with the physical system to be simulated. It is acknowledged as similar to any physical models or correlations in the code: i.e. system nodalization presents an inherent code uncertainty and like code models and correlations, quantification of nodalization-based code uncertainty is deemed to be equally important to quantify model accuracy and uncertainty.

For application to nuclear reactor safety relevant issues, i.e. FSAR Chapter 15 transient analyses, two broad groups of nodalizations shall be distinguished when applying a BE code-nodalization to the analysis of nuclear reactor safety relevant issues. They are briefly recalled in the sequel and fully described in the SAPIUM report:

a) The code focus is a component or a 'simple' system having parameters directly recognized by the code structure (no, or limited user interpretation is needed): in this case the effort requested to the code user is to implement the system or the component properties (typically geometric and thermodynamic) into the nodalization with limited or no need for engineering judgment;

b) The code has a modular nature and the system to be modeled is complex. In this case, expertise is requested to the user 'to make readable' the system peculiarities to the code. This is achieved throughout the nodalization whose development does require engineering judgment.

The distinction between categories a) and b) is relevant to the present framework. In the first category, the nodalization can be presumed as embedded into the code architecture and the code validation implies also the validation of the nodalization approach. In the second category, the system nodalization is developed by dividing the real plant/facility component volumes into a set of control volumes that are essentially stream-tubes having inlet and outlet flow path connections. Subdivision of such a complex system can be done in a number of ways but for a successful solution in the case of the analysis, a number of factors must be satisfied: numerical stability, run time, and spatial convergence. In addition, engineering judgment is normally used to a wide extent to develop the system nodalization.

Experience with code assessment case studies and International Standard Problems have shown that the nodalization structure is subjected to the 'code user-effect' to quite a large extent and the nodalization itself constitutes an "independent" computational tool which shall undergo through an assessment of applicability. The importance of establishing a procedure for the nodalization set-up and the assessment of its applicability as part of the Evaluation Model (EM) (Petruzzi et al., 2019) is a consequence of the above mentioned complexity and subjectivity of the process.

2.4.2.2. Assessment of Applicability of the Simulation Model

The process for assessing the applicability of the SM has the goal to demonstrate that the SM calculation results - obtained by the application of the best estimate code with the nodalization developed following the standard nodalization rules and techniques - constitute a realistic approximation of the reference behavior of a test facility (Petruzzi et al., 2019; Petruzzi and D'Auria, 2016; D'Auria et al., 1995) . The process should take into account the effect of many different sources of approximations:

- The data of the reference test facility available to the code-user are typically non exhaustive to reproduce a perfect nodalization of the reference test facility;
- The code-user derives, from the available data, an approximated nodalization-schematization of the facility reducing the level of details of the simulated hardware;
- The code capability to reproduce the hardware, the systems (test facility or plant) and the actuation logic of the systems further reduce the level of detail of the nodalization-schematization.

The development of the process for assessing the applicability of the SM should include:

- the establishment of necessary requirements for the different aspects above mentioned, and
- the determination of whether or not those requirements are met by the SM for all selected experiments in the database.

The goal is to develop and obtain a qualified SM considering the comparison with the hardware data, the BIC and the time trends of relevant quantities. The process should distinguish at least between three main sub-steps:

1) demonstration of the geometrical fidelity of the nodalization,
2) demonstration of the achievement of the steady state, qualitative transient analysis and
3) quantitative accuracy evaluation.

Criteria for selecting relevant quantities in each of the above three sub-steps should be defined as well as the definition of the assessment applicability requirements. If any requirement in any of the three sub-steps is not fulfilled, the process of assessment of applicability of the SM is not passed and the main elements of the SM should be improved. A new process should be then applied in full, i.e. all three sub-steps, to the modified SM. It should be also emphasized that the three-steps process for assessing the applicability of the SM apply to all tests of the experimental database used for the input uncertainty quantification and validation and if the process fails for any of the test, the SM should be modified and improved and a new process should be applied to the modified SM for all tests of the experimental database.

2.4.2.3. Construction and verification of consistency indicators (accuracy code/experiment) and definition of a scale of accuracy for qualification of the simulation model

A consistency indicator is the basis for comparing responses from experimental data with simulation model predictions. The primary consideration for establishing a consistency indicator should be what the model must predict in conjunction with what types of data available from the experiment. Additionally, the indicators should provide a measure of agreement that also includes estimates of the numerical and experimental errors (Leonardi et al., 1994; Cacuci, 2019).

The SAPIUM report provides the description of different types of indicators as well as the requirements for their construction. The values of these indicators are then compared to selected thresholds-values requirements.

In carrying-out the assessment of adequacy of the SM, a range of tests (with different boundary and initial conditions and at different scale) should be employed to demonstrate that the SM has not been tuned to a single test. For integral behavior assessment, counterpart tests (similar scenarios and transient conditions) in different experimental facilities at different scales should be selected. Assessments using such tests lead to information concerning scale effects on the models used for a particular SM.

2.4.3. Step 9 - Selection of important uncertain input parameters by sensitivity analysis

The models of complex physical systems, like the BE SM for conducting the safety analysis of NPP, are law-driven models characterized by the presence of balance equations and several correlations. As a consequence, those models are customarily over-parametrized (e.g. thousands of input parameters are needed to build a typical SM for NPP), as they may include more relevant laws than the amount of available data would support for the validation. For the same reason, those models may have also a greater capacity to describe the system under unobserved circumstances (outside the ranges of derivation of the correlations – see the list of 'sources of uncertainty' – i.e. outside the validation domain), even though the crucial question stays on the uncertainty of those predictions respect to the reality.

When focusing on the selection of input parameters, it is important to make a clear distinction between 'important' parameters as those whose uncertainty contributes substantially to the uncertainty of the output results, and 'sensitive' parameters as those which have a significant influence on the output results. This distinction is in the type of analysis being conducted: global (that addresses parameter importance) and local (that addresses the parameter sensitivity) sensitivity analysis.

In practice, the sensitivity analysis (SA) process is an invaluable tool (Sobol, 1993; Saltelli et al., 2000). At the end of a sensitivity analysis, the analysts will hold a 'sensitivity ranking' of the input parameters sorted by the amount of influence each has on the model output. Disagreement among rankings by the various SA methods for parameters less sensitive is not of practical concern since these variables have little or no influence on model output. This last statement - 'little or no influence on model output' - might be false or completing misleading when an importance analysis has to be considered: in this situation given the practical difficulties/impossibilities to identify input uncertainties, the conclusions that can be drawn from the actual ranking might not be appropriate to estimate the uncertainty of the output model responses (it is for instance the case of a very low sensitive input parameter with a large range of uncertainty).

2.5. Element 4 - model input uncertainty quantification

Exploiting the experimental database and the simulation model fully characterized in the previous sections, this fourth element is devoted to the uncertainty quantification of model input parameters, which is the so-called inverse propagation problem (IP). This has been treated in the previous PREMIUM project (Mendizábal et al., 2017). The four following steps (Figure 5) are identified in this element.

2.5.1. Step 10 - Aggregation of the information coming from experiments and simulation model to be used in the "inverse propagation"

This step consists in aggregating the information coming from Elements 2 and 3 before performing the IUQ. In practice, it appears that different pairs of simulated/experimental value can have different degrees of importance for the intended use. This degree of importance depends on the representativeness of each experiment of the database (and a clear analysis is needed in Element 2) as well as on the type of SRQs considered. This should be integrated in the development of the quantification method to avoid under or overweighting the influence of a given pair.

Figure 5. Steps of Element 4.

2.5.2. Step 11 - Quantification of model input uncertainties by "Inverse propagation"

Many inverse propagation (IP) methods (both for calibration and uncertainty quantification of models) already exist in the nuclear safety field, and more are presently under development. The OECD/NEA PREMIUM project was devoted to application and comparison of inverse methods for quantification of model uncertainties on the basis of CETs. The benchmark was performed using reflood experiments (Mendizábal et al., 2017) . A review and comparison of the available IP methods used in PREMIUM is given in (Reventós et al., 2016) .

The simplest and the most widespread method of invers quantification of model uncertainties is the quantification on the basis of SETs, where singular measurement representing the phenomenon can be compared with associated calculated parameter (Skorek, 2004), e.g. measured void fraction in the channel is associated with phase relative velocity, only. The population of point values (multipliers) obtained from the comparison can be approximated as Probability Density Function (PDF) representing model uncertainty, in this care relative velocity model. Among the advanced methods addressed particularly for CETs, the most usual modelling of uncertainty is the probabilistic one. To solve a probabilistic IP is an exercise of statistical inference, and two basic approaches exist: frequentist and Bayesian (Mendizábal, 2018). The first one is based on Maximum Likelihood inference (Kuhn, 2003). The second one exploits the Bayes theorem (Gosh et al., 2006), which is a procedure of updating information. There exist also in the literature alternative approaches combining the use of design of experiments and forward uncertainty propagation though they are much less prevalent than probabilistic ones (see (Kovtonyuk et al., 2015) for an example in the nuclear field, and (Freixa et al., 2016) for a comparison exercise with a methods based on Maximum Likelihood inference). The SAPIUM report provides a full description of available methods.

The choice of an inverse method depends on the problem under study, and should be made according to the following criteria:

- Solidity: rigorous mathematical framework.
- Flexibility: different situations associated to the problem can be taken into account. E.g. experimental uncertainty, model bias estimation…
- Transparency and reproducibility: clear documentation available, including list of assumptions and user guidelines. Uncertainty associated to the method is an important information.
- Reduction of user effect / expert judgment.
- Relevance of method assumptions for the problem under study. E.g. checking if the nature of uncertainties (aleatory and epistemic) and the lack of information on them are properly taken into account.
- Tractability: e.g. limitations due to high computational cost can be circumvented through the use of surrogate models or efficient design of experiments.

2.5.3. Step 12 - Combination of model input uncertainties if several quantifications are performed

This step concerns the development of procedures for combining the results of different quantifications . This is not always required in practice, since a first aggregation is achieved in Step 10. But in Step 10 the aggregation of information is previous to the uncertainty quantification, while in Step 12 it is posterior to it. The different quantifications can arise from different studies, or to the different groupings of experiments (e.g. according to their scale).
Synthesizing the information implies also evaluating possible conflicts (i.e. disagreement) among the different quantifications. A strong conflict would require some iteration of the SAPIUM steps-elements.
Diverse techniques of information synthesis can be applied in this step. For instance, weighted averages of probability distributions can be constructed. Bayesian techniques can be applied (requiring specification of prior distributions). There are also techniques based on possibility theory (Dubois et al., 2016), defining different types of "fusion operators" (conjunctive, disjunctive, arithmetic mean), and allowing also the definition of conflict or disagreement indicators.

2.5.4. Step 13 - Confirmation by counterpart tests

Once the model input uncertainty quantification has been performed, a confirmation step is needed, consisting of making the forward propagation of the obtained model input uncertainty for counterpart tests, and checking if the response data used in the quantification are adequately enveloped.

From a technical point of view, confirmation is similar to validation. It especially requires the use of mathematical indicators to quantify the agreement between simulation and experimental results. This topic is addressed in the next element of the SAPIUM process.

2.6. Element 5 - model input uncertainty validation

Following the VVUQ formal procedure (Ferson et al., 2008; Oberkampf and Barone, 2006), the technical treatment of the validation process encompasses firstly a comparison between the simulation model output uncertainty and experimental data not used in the quantification and secondly a predictive capability or adequacy assessment outside the experimental domain. This last task is discussed in the conclusion of this paper and the following 4 steps (Figure 6) are introduced to perform the first task.

Figure 6. Steps of Element 5.

2.6.1. Step 14 - Determination of numerical approximation and other input data uncertainties for each validation case

In the case of input uncertainties, the validation cannot be done in the input space since the comparison of the results with experimental data is not possible. It is performed in the SRQ space after input uncertainty propagation through the simulation model. Therefore, it is mandatory to combine the quantified model input uncertainties with other sources of input uncertainties (e.g. due to model form and numerical approximation) because the validation result is also affected by them. The SAPIUM activity is focused on model input uncertainty, therefore, the quantification of the other types of uncertainties are assumed to have been performed in another part of the VVUQ process.

2.6.2. Step 15 - Propagation of all input uncertainties through the simulation model

Among all the available uncertainty analysis methods, the input uncertainty propagation approach is very popular in industrial applications. It has been used in the framework of nuclear safety analysis in (OECD/NEA, 2011) for example. It combines mathematical modeling of input uncertainties and Monte

Carlo techniques to estimate statistical quantity of interest. The simulation model is treated as a “black box”, and the input uncertainties are propagated to the SRQ uncertainties via several simulation model runs.

2.6.3. Step 16 - Computation of validation indicators

The comparison between simulation (after input propagation) and experimental results then exploits validation indicators. A straightforward comparison can consist in checking if the experimental value falls inside the SRQ uncertainty interval. In this case, a very large uncertainty band can be considered as acceptable since it is more likely to encompass the experimental value. Therefore, a fully transparent construction of validation indicators should rely on an appropriate definition of the target quantity of validation (interval, cumulative distribution function,…) and of the important characteristics of the SRQ uncertainty to capture for validation. This involves the use of more complex validation indicators to have a better qualitative insight on the results.
Most of the classical validation indicators in the literature (Liu et al., 2011; Oberkampf and Barone, 2006) are referred as calibration ones i.e. they focus on a consistency checking between the information obtained after propagation of quantified input uncertainties and experimental values. Among them, one can mention the so-called hypothesis testing approach or area metric (Ferson et al., 2008). In some situations, it can happen that a satisfactory calibration is reached thanks to a very wide uncertainty band that is more likely to encompass an experimental value. However, an artificially large uncertainty could be difficult to analyze. The characteristics (and especially the width) of the uncertainty band should therefore be considered. Distinguishing concentration of the information associated to the SRQ uncertainty and calibration has been already addressed in many scientific fields (Cooke, 1991; Destercke and Chojnacki, 2008; Gneiting et al., 2007; JCGM, 2012).

2.6.4. Step 17 - Analysis of the validation results

Finally, the validation results are analyzed. For each SRQ, the evaluation of the agreement between simulation and experimental results requires the definition of a scale of acceptability. This scale obviously depends on the type of validation indicator. If this indicator only consists in checking if each experimental value falls inside each corresponding uncertainty band, it should be based on the percentage of experiments for which this previous agreement check is satisfied. When considering a more complex indicators such as hypothesis testing, significance thresholds are available but they rely on assumptions (e.g. independence of experiments) and a careful check of their fulfillment has to be performed.

Moreover, keeping in mind that the objective is application to NPPs, the analysis of the results provided by the computation of the validation indicators is not sufficient. Adapting the works of (Hemez et al., 2010) or (Oberkampf et al., 2007) to the IUQ framework, this analysis in the validation domain should be combined with the adequacy of the experimental database. It leads to the evaluation of the maturity of the IUQ for the intended use. This evaluation should be carried out in a loop approach. If the maturity is proved as not sufficient, an iteration step is performed. It means the maturity is improved according to findings during quantification and validation and the whole procedure is repeated. A lack of maturity can be due to different choices adopted in each SAPIUM element. One reason could be the lack of adequacy of the experimental database. In this case, it is preferable to consider integral experiments in the quantification (and not only in the validation) and to iteratively revise the quantified model input uncertainties from these new data.

2.7. Scaling issues and predictive capability assessment

The main concern of the predictive capability of uncertainty analyses is geometrical and thermal-hydraulic scaling. It means, if the identification and quantification of model uncertainties on the basis of selected experiments are valid for large scale geometry and range of thermal-hydraulic parameters characteristic for nuclear reactors. The problem of the up-scaling in the best estimate thermal-hydraulic simulations is a central problem in the nuclear reactor safety. This is a general problem and affect all activities in this field: experimental work as well as numerical analyses (Roy and Oberkampf, 2011). This topic was a subject of intensive investigations and review of the activities and findings in this field was analysed in the frame of OECD/NEA projects (Bestion et al., 2017).

The dependency of interesting phenomena according to change of thermal-hydraulic conditions like pressure or temperature are usually known. It is a common practice to investigate the dependency of phenomena on thermal-hydraulic parameters for the full range of possible applications. The geometry of the experiments is mostly small-scale (or medium-scale) and rather only exceptionally equivalent to full-scale nuclear reactors. So, the problem of the up-scaling appears to be the main concern of the uncertainty analysis. Investigations performed in the past showed that the findings obtained on the basis of small-scale experiments are applicable only partially for large scales.

The conclusion was that the scale-up effects are to be considered by the development of a qualified input data set for each reference (best estimate) calculation and by selection and quantification of uncertain input parameters. In particular, differences in uncertainties of physical models according to their application to different scale objects have to be taken into account.

Other possibility to consider scaling effects is selection of different correlations according to their field of application. If there is such option in the physical model of the thermal-hydraulic code, different correlations/constitutive equations may be applied for small and large scale facilities according to the recommendation in the code documentation. Such recommendations result from code development and validation and as such express the state of knowledge concerning also the scaling effect.

The ranges variation of physical model uncertainties is the main way of scale-up effect consideration in uncertainty analyses. Since the quantification of model uncertainties takes place by comparison with experimental data, an appropriate selection of the adequate experiments is of importance. The preferable model uncertainties quantification is comparison of code predictions with experimental data from SETs. The experimental data selected for quantification have to be representative for the considered application. In particular they have to reflect the scale of the analysed facility. The optimal situation is, when there are available SETs for the whole spectrum of scales where the model will be applied. The experiments considered by evaluation of scale effect underlay the same general requirements regarding adequacy to investigated phenomena and geometry and accuracy of experimental measurements, applied for selection of tests for experimental data base as described in the Section 2.3.

The condition for a correct consideration of scale-up effects is to carry out carefully complete uncertainty and sensitivity analyses for each application. The results of uncertainty analyses for small scale facilities are important source of information and experience but cannot be directly transformed to large scale application. The most important step by consideration of the scale-up effects is the identification and quantification of input uncertainties, in particular model uncertainties for large scale applications. Since some large scale separate effect experiments exist, the quantification can be performed in the best way on the basis of comparison with available experimental data. Once the model uncertainties have been quantified the propagation of the input uncertainties through the mechanistic codes enables carrying out of the best estimate plus uncertainty analyses for any transient or accident in the field of the code application; also for events for which integral tests do not exist. This capability is a clear advantage of the uncertainty estimation method based on input uncertainties propagation (e.g. using Wilks' formula (Wilks, 1941)).

However, it requires a proper quantification of input uncertainties and sufficient experimental basis of SETs or CETs for model input uncertainties quantification.

In the case of lack of suitable experiments, the new trend is using CFD simulations as basis for evaluation of system codes models, e.g. (Lewis et al., 2016). Application of CFD simulations instead of experimental data for uncertainties evaluation of 1-D system code models is a relatively new option but of increasing importance. However, this approach first requires performing a VVUQ analysis of the CFD simulation model.

Another aspect of the prediction capability is related to limitation of IUQ performed for a particular transient (specific approach contrary to generic approach). It can happen that in the course of uncertainty analysis the range of varied calculations extends the assumed range of parameters considered for development of experimental data base and following model input uncertainties quantification. In such a case the frequently recommended best solution is performing of iteration step. Beginning with extension of the experimental data base, quantification and finally validation of model uncertainties have to be performed once more.

The predictive capability and more general safety of nuclear reactors relay on quality of extrapolation the existing information to the full-scale application of nuclear reactors. As an ultimate measure of ensuring security of the safety analysis of nuclear power reactors, lack of exact information on uncertainty of the upscaling to the full-scale NPP applications in nuclear reactor safety is compensated by application of safety margins.

3. Main recommendations

This section summarizes the main lesson learned from the analysis of the SAPIUM elements.

3.1. Element 1- specification of the problem and requirements

It is recommended to clearly specify the model IUQ problem for the developed or selected simulation model, according to the transient classification as documented in the FSAR of the targeted NPP (Wilson and Boyack, 1998), or in case of new plant design, to specify the application domain for the newly identified transient scenarios. It is recommended to choose the SRQs based on parameters that are directly and accurately measured in the experiments, which are used for verification of the design limits or surrogates in the NPP accident analysis (ANSI, 1973, 1988). It is recommended to group different accident scenarios (e.g., RCS heatup or cooldown accidents, reactivity initiated accidents) (USNRC, 1978) into a single IUQ problem with common SRQs of interest, and make a generic model input uncertainty quantification and validation for the developed or selected simulation model.

The NPP accident scenario identification and IUQ definition process can rely heavily on expert opinion and can be subjective. Therefore, iteration of the process, based on experimentation and analysis, is important. It is recommended to use the phenomenon identification and raking table (PIRT) technique (Diamond, 2006; Wilson and Boyack, 1998) to first identify and rank the physical phenomena, and use the sampling-based global sensitivity analysis (GSA) technique to confirm the PIRT (also called Q-PIRT) (Luo et al., 2010; Martin, 2011; Yurko and Jacopo, 2012).

It is important to keep in mind the phenomena ranking could be subject to the limitations of the knowledge of the expert, the experimental databases or simulation models. An iteration with other elements may be necessary if such limitations are identified during the SAPIUM process.

3.2. Element 2 - development and assessment of the experimental database

It is recommended to follow a structured and transparent approach to perform the construction of a scenario-dependent experimental database.

It first requires describing in a standard format (including information on covered phenomena, geometry, scaling effect, experimental uncertainties) all available experiments coming from SETs, IETs and CETs. Various scales IETs are mandatory for the application to the reactor case. Extra experiments might be also required if the adequacy of the database is not sufficient.

For a transparent and reproducible adequacy assessment, it is recommended to use mathematical tools to quantitatively perform the analysis. Multi-Criteria Decision Making approach can be exploited in order to objectively and automatically evaluate the representativeness leading to a ranking of experiments. The evaluation of the completeness of the experimental database is based on a completeness index. This index should be restricted to the database and not be applied to the whole quantification/validation process as it is classically done in VVUQ.

Finally, a special attention should be devoted to the splitting of the experimental database for input uncertainty quantification and validation. If the number of available experiments is too limited to perform this splitting, all available experiments should be considered for the quantification step and the validation step should be adapted.

The SAPIUM approach is proposed to improve the IUQ by expert judgement. However, in some situations the expert judgement could still be used. A typical situation is when available measured parameters are not sufficient to determine separate uncertainties. This is usually the case by integral experiments. This problem has to be treated by quantification and validation step. Moreover, expert judgement might be also required in case of indirect link between measures and input parameters. In such situations, the problem needs to be solved by quantification procedure.

3.3. Element 3 - selection and assessment of the simulation model

The assessment of the applicability of the SM is an iterative process that applies to all experimental tests of the validation database. The possible failure of the process for one test implies the improvement of the SM (either code or nodalization or both) and the repetition of the assessment of the applicability of an SM for all tests of the validation database.

The validity of a simulation model is defined over the domain of model form, inputs, parameters, and responses. This fact effectively limits use of the model to the particular application for which it was validated; use for any other purpose would require the assessment of the applicability of the SM to be performed again. In other words, the assessment of the applicability process cannot prove that an SM is correct and accurate for all possible conditions and applications, but, rather, it can provide evidence that a SM is sufficiently accurate. Therefore, the assessment of the applicability process is completed when sufficiency is reached.

The SM shall not be tuned to a particular data set and the data used to assess the SM have not been deliberately selected to make the SM appear to be more accurate than it truly is.

An important aspect of the simulation model assessment is the nodalization strategy and model option selection. They should be consistent between the experiment and the nuclear power plant.

In the assessment process, a special attention should be devoted to the construction of consistency indicators to evaluate the accuracy between simulation and experiment.

The selection of important uncertain input parameters (including nature of uncertainties e.g. aleatory, epistemic) should be confirmed by sensitivity analysis methods, in order to reduce the subjectivity by expert judgement.

3.4. Element 4 - model input uncertainty quantification

Inverse methods can also be used to calibrate a model (simultaneously or not with quantification). The experience in PREMIUM indicates that the IUQ should be performed without recalibration. This is especially advisable in case of lack of adequacy of the available experimental database.

Inverse methods should allow assigning different weights to different pairs of simulated/experimental value in order to avoid under or overweighting the influence of a given one. It is the case for example for probabilistic ones that use the so-called likelihood function in their inferences, and a possibility is to construct a weighted likelihood. The least-square technique, used in the solution of inverse problems, can also assign different weights in the construction of the sum of squares to be minimized. In any cases, it is compulsory to perform sensitivity analysis to the assigned weights in order to ensure the reliability of the quantification.

The assumptions associated to the inverse problem mathematical methods (type of method, uncertainty modeling to handle aleatory and epistemic uncertainties, etc) should be clearly taken into account to evaluate the impact of the analyst's choices on the results. Moreover, depending on the problem to solve, users may choose between probabilistic or non-probabilistic modeling (i.e. based on alternative uncertainty theories). For instance, an inverse problem where the noise has a known probability distribution (e.g. Gaussian) can benefit from probabilistic methods. In presence of incomplete knowledge on input uncertainties, it is advisable to combine different methods to avoid formulating extra assumptions.

More generally, the choice of an inverse method depends on the problem under study, and should be made according to the main criteria introduced in Section 2.5.2.

Influence of additional uncertainty sources (e.g. related to the numerical approximation and especially to the choice of the nodalization) is of prime importance to derive reliable input uncertainties. The development of inverse methods allowing the combination of different uncertainty sources is therefore advisable.

It is important to remember that inverse problems, in general, do not have unique solutions, and additional information must be added in order to select one of the possible solutions. For example, different users can differ in the way of introducing a given prior information, thus obtaining different results.

3.5. Element 5 - model input uncertainty validation

The main recommendations can be split with respect to 4 different topics.

The first one concerns the validation experiments. If the experimental database is large enough, they should correspond to experiments which are not used for the uncertainty quantification. When the number of experiments is not sufficient to split the database, a leave-one-out cross-validation which is a classical technique in statistics, e.g. (Wackernagel, 1998), can be exploited. Another option could be to enforce the connections between experimentalists and developers to design new validation experiments following specific requirements (Oberkampf and Trucano, 2007).

It is also important to control (or evaluate) the impact of the methodological assumptions on the validation result. These assumptions are first related to the input uncertainty modelling that should integrate the state of knowledge on uncertainties (e.g. aleatory/epistemic uncertainties). They also concern the construction of validation indicators that can rely on strong assumptions on experiments (e.g. independence of experimental conditions) that are not always satisfied in practice. This should be taken into account when the acceptability of uncertainty results is checked for the intended use by comparison to acceptability thresholds.

The third topic is related to the selection of a validation indicator. It is recommended to clearly state the choice of the target quantity to validate defining the SRQ uncertainty and the important characteristics of this uncertainty to capture for validation before starting the study. Besides a consistency checking, it is advisable to take into account extra features such as the concentration of the information provided by the uncertainty analysis by combining informativeness and calibration type indicators.

Finally, the last topic is focused on the analysis of the validation results. The computation of the validation indicators is not sufficient if the objective is application to NPPs. It is recommended to construct a maturity model combining the acceptability of the validation results in the validation domain with the adequacy of the experimental database. The evaluation of the predictive maturity should be carried out in a loop approach. If the lack of maturity can be explained by a lack of adequacy of the experimental database, it is preferable to extend the database by considering integral experiments for the quantification and to iteratively revise the quantified model input uncertainties from these new data.

3.6. Scaling issues and predictive capability assessment

Since some large scale separate effect experiments exist, the quantification should be performed in the best way on the basis of comparison with available experimental data, as far as possible.

Concerning the quantification of model input uncertainties on the basis of SETs, the methodology is matured and widely used. In the last time a number of methods for simultaneous quantification of several model input uncertainties on the basis of CETs have been developed, e. g. (Wu et al., 2018), (Liu et al., 2018) and successfully applied. The main problems are the semi-empirical closure relations and their uncertainty prediction for the small and large scale geometry. The results of uncertainty analyses for small scale facilities are important source of information and experience but cannot be directly transformed to large scale application. In the case of lack of experimental evidence for large scales, the extrapolation of quantified model uncertainties has to be performed carefully and increase of the uncertainty ranges has to be taken into account.

An optimal way of obtaining reliable model input uncertainties would be evaluation of the physical model uncertainties by the code/model developers exploiting advanced validation (Unal et al., 2011) or extended validation (Skorek, 2017). However, complete quantification of all relevant models in the code application field is a difficult issue which requires a lot of resources. Even, if there is an obvious interest of performing systematic evaluation of code models uncertainties, it would require still a lot of work and a long time to complete the process.

Once the model input uncertainties have been quantified the propagation of the input uncertainties (e.g. using Wilks' formula) through the simulation model enables carrying out of the best estimate plus uncertainty analyses for any transient or accident in the field of the code application; also, for events for which integral tests do not exist. However, it requires a proper quantification of input uncertainties and sufficient experimental basis of separate effect tests for model input uncertainties quantification.

A clear deficiency of performing uncertainty analyses for reactor scale geometries is lack of large scale experiments. CFD calculations can be used in some cases, mainly for single phase flows, as reference for system code simulations. However, in the case of two-phase flow the accuracy of CFD predictions is still limited. Therefore, additional large-scale experiments are of importance and would be a great help for quantification as well as validation of model uncertainties.

4. Open issues for future development

The work of the SAPIUM group has revealed several remaining open issues for a complete application of the proposed IUQ approach. They concern:

4.1. Adequacy of the experimental database

IUQ process is based on the comparison between simulation and experimental results. Therefore, the quantified model input uncertainties strongly depend on the adequacy of the experimental database. Two main aspects of this topic should be further investigated.

The first one is related to the quantitative analysis of a database and to the construction of representativeness and completeness indices. Several generic tools are already developed but their extension to the framework of IUQ still remains an open issue.

The second one concerns the lack of experiments. In the case of a poor validation, the SAPIUM approach provides some tools and recommendations to evaluate if the experimental database should be enlarged (e.g. computation of adequacy indicator, loop approach, cross-validation). However, in practice, it is not always affordable when the number of available experiments is too limited and that the few available experiments do not define an adequate experimental database. There is no clear strategy to deal with this last situation and to measure the impact on the acceptability of the quantified model input uncertainties.

4.2. Integration of the state of knowledge in the construction of inverse methods

There is a large literature related to the construction of inverse methods. However, the degree of reliability (for the problem under study) of the information provided by each pair experimental/simulation value coming from Elements 2 and 3 of the SAPIUM approach is hardly taken into account. It requires the construction of procedures to combine information with different degrees of importance.

Moreover, most of the available inverse methods are constructed in the probabilistic framework and requires the choice of unique pdfs to model uncertainties. In case of incomplete knowledge, the use of alternative uncertainty model could be an interesting strategy. If several works have been already proposed to treat the direct problem, adaption of probabilistic inverse methods to alternative theory remains a challenging problem in the framework of nuclear applications. In particular, further investigations are required to tackle the problem of treatment of epistemic uncertainty by alternative theories for IUQ (possibility, Dempster-Shafer,…).

4.3. Acceptability of the validation results

Important efforts should be also devoted to the analysis of the validation results in the validation domain with the objective of extrapolation to the application one. It first involves the construction of a reliable model for the evaluation of the predictive maturity of the whole process. A general guideline has been provided in the VVUQ framework as well as a scale of maturity but adaption to IUQ and combination with adequacy indices still remain open questions.

The acceptability of the validation results strongly depends on the quantification of other uncertainty sources (numerical approximation error and model form) in the validation process. The contributions of the SAPIUM document are restricted to quantification of model input uncertainties. However, the validation, since it operates in the SRQ space, will be affected by all the types of uncertainty sources. A poor validation that would require iterating the SAPIUM process can therefore be due to other uncertainty sources than model input uncertainties. To avoid any misinterpretation, further works are necessary to understand and evaluate the impact of each category of uncertainty sources on the validation results.

4.4. Predictive assessment and extrapolation to the application domain

The main weakness of the predictive capability of application inverted uncertainties quantification in the field of Reactor BEPU analyses is lack of suitable large-scale experiments. As long as there is available an adequate experimental basis and quantification of the input uncertainties can be proved by validation procedure, the predictive capability for large scales applications is ensured.
But frequently it is not the case, available are only small-scale experiments, and the quantified uncertainties needed to be extrapolated for the analyses in the reactor scale. There were performed some efforts to solve this problem by development of integral methods for estimation of scale effects in BEPU analyses, e.g. EMDAP methodology (Dzodzo, 2018). A possible way is, basing on scaled integral test, extrapolation of behaviour of reactor systems from test facilities to the reactor safety analyses using system codes as extrapolation tools (Mascari, 2015). But there is no matured extrapolation method, which could be recommended without a doubt for such applications.
This is the main issue to be solved. May be an option for the farther investigation in this field could be combination of input uncertainties propagation method with methods based on output uncertainties extrapolation.

5. Conclusions

This paper was devoted to the description of the different elements and steps of a good practice guidance for a transparent and rigorous model IUQ. The main developments have been performed by the SAPIUM group that included contributors from 10 organizations including Technical Support Organization (TSO), industry and university. The starting point was the available state of knowledge coming from previous related OECD/NEA projects as well as current practices in regulation, industries and research. The main outcome is an original structured approach that offers a general framework to develop or analyze IUQ.

This work clearly emphasized that IUQ shouldn't be reduced to the application of inverse methods as it was the case during the PREMIUM activity. It is a more general process that involves a clear specification of the problem and efficient strategies to construct adequate experimental database and to combine the information coming from different experiments. It also requires assessing the simulation model before quantifying input uncertainties. Finally, the validation of the quantified model input uncertainties has to be taken into account in the whole process in order to check the acceptability of the results for the intended use. The description of the different elements revealed the need to use mathematical tools integrating the physical knowledge in order to be fully rigorous, transparent and reproducible.

Even if part of the material to perform model IUQ is available, several remaining open issues should be tackled. This paper includes some development axes for future works. They concern

- the adequacy analysis of the experimental database: construction of quantitative indices to evaluate the adequacy, development of a strategy for a limited number of experiments,
- the construction of inverse propagation methods: integration of a reliability degree of the information to back-propagate, treatment of aleatory and epistemic uncertainties by uncertainty theories,

- the evaluation of the acceptability of the results: construction of predictive maturity indices devoted to IUQ, evaluation of the impact of other uncertainty sources on the validation, construction of a matured extrapolation process to exploit the quantified input uncertainties in the application domain.

For industrial application, it is important to focus on the practicability of the proposed tools. This can be achieved by exercises of comparison and benchmark of IUQ constructed in the SAPIUM framework. As a SAPIUM follow-on activity, demonstration cases could be conducted, with the objectives to verify (i) the applicability of the best-practices, (ii) that whatever the problematic, the best-practices allow to deal with, and (iii) the issues identified during PREMIUM can be avoided by following the SAPIUM best-practices. For that, these demonstration cases could be split into the following three steps:

- 1st step: The experimental database contains enough data and there are few influential phenomena: the problem is well posed and seems easy to treat.
- 2nd step: In this case, several influential phenomena are involved, but the experimental database provides enough data of different types. An example of this case is the critical flow at the break, which is of significant relevance during a LOCA.
- 3rd step: In the last case, several influential phenomena are present, but the experimental database does not contain enough varied data (such as in PREMIUM benchmark).

This progressive approach could help to treat the key issues identified and to use gradually the recommendations developed in SAPIUM, and give a feedback on the proposed methodology.

NOMENCLATURE

AHP	Analytical Hierarchical Process
BIC	Boundary and Initial Conditions
BE	Best-Estimate
BEPU	Best-Estimate Plus Uncertainty
BWR	Boiling Water Reactor
CANDU	CANada Deuterium Uranium (reactor)
CCVM	CSNI Code Validation Matrix
CET	Combined Effect Tests
CFD	Computational Fluid Dynamics
CSNI	Committee on the Safety of Nuclear Installations
EM	Evaluation Model
EMDAP	Evaluation Model Development and Assessment Process
FSAR	Final Safety Analysis Report
GSA	Global Sensitivity Analysis
IET	Integral effects Test
IP	Inverse Propagation
IUQ	Input Uncertainty Quantification
LOCA	Loss Of Coolant Accident
LWR	Light-Water Reactor
MCDM	Multi-Criteria Decision Making
NEA	Nuclear Energy Agency
NPP	Nuclear Power Plant
OECD	Organization for Economic Co-operation and Development
PDF	Probability Density Function
PIRT	Phenomena Identification and Ranking Table
PMI	Predictive Maturity Indexes
PREMIUM	Post-BEMUSE Reflood Models Input Uncertainty Methods

PWR	Pressurized Water Reactor
Q-PIRT	Quantitative Phenomena Identification and Ranking Table
RCS	Reactor Coolant System
SA	Sensitivity Analysis
SAPIUM	Systematic APproach for model Input Uncertainty quantification Methodology
SET	Separate Effects Test
SM	Simulation Model
SRQ	System Responses Quantity
TSO	Technical Support Organization
USNRC	United States Nuclear Regulatory Commission
VVUQ	Validation and Verification, Uncertainty Quantification

Acknowledgments: The authors are grateful to OECD/NEA for organizing the SAPIUM project, and to their respective organizations for supporting their participation in this project.

.