



Quantification of the uncertainty of the physical models in the system thermal-hydraulic codes – PREMIUM benchmark

Tomasz Skorek^{a,*}, Agnès de Crécy^b, Andriy Kovtonyuk^{c,m}, Alessandro Petrucci^{c,q}, Rafael Mendizábal^d, Elsa de Alfonso^e, Francesc Reventós^e, Jordi Freixa^e, Christine Sarrette^f, Milos Kyncl^g, Rostislav Pernica^g, Jean Baccou^h, Fabrice Fouet^h, Pierre Probst^h, Bub-Dong Chungⁱ, Tran Thanh Tramⁱ, Deog-Yeon Oh^j, Alexey Gusev^k, Alexander Falkov^k, Yuri Shvestov^k, Dong Li^l, Xiaojing Liu^l, Jinzhao Zhang^m, Torsti Alkuⁿ, Joonas Kurkiⁿ, Wadim Jäger^o, Victor Sánchez^o, Damar Wicaksono^p, Omar Zerkak^p, Andreas Pautz^p

^a Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) GmbH, 85748 Garching, Germany

^b Commissariat à l'Energie Atomique (CEA), CEA-Grenoble, 38054 Grenoble Cedex 9, France

^c University of Pisa, San Piero a Grado Nuclear Research Group, 56122 San Piero a Grado, Pisa, Italy

^d Consejo de Seguridad Nuclear (CSN), 28040 Madrid, Spain

^e Universitat Politècnica de Catalunya (UPC), Avda. Diagonal 647, 08028 Barcelona, Spain

^f Bel V, 1070 Brussels, Belgium

^g Research Centre Rez Ltd. TSO (CVRez), Rez, Czech Republic

^h Institut de Radioprotection et de Sûreté Nucléaire (IRSN), 13115 St Paul-Lez-Durance, France

ⁱ Korea Atomic Energy Research Institute (KAERI), Daejeon 34057, Republic of Korea

^j Korea Institute of Nuclear Safety (KINS), Daejeon 34142, Republic of Korea

^k OKB Mechanical Engineering (OKBM), Nizny Novgorod, Russia

^l Shanghai Jiao Tong University (SJTU), Shanghai, China

^m Tractebel (ENGIE), Boulevard Simon Bolivar 34-36, 1000 Brussels, Belgium

ⁿ VTT Technical Research Centre of Finland, Espoo FI-02044 VTT, Finland

^o Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

^p Paul Scherrer Institute (PSI), 5232 Villigen PSI, Switzerland

^q NINE, Nuclear and Industrial Engineering, Lucca 55100, Italy

ARTICLE INFO

Acronyms

BEPU
Best Estimate plus Uncertainty
CET
Combined Effect Tests
CFD
Computational Fluid Dynamic
CIRCÉ
Calcul des Incertitudes Relatives aux Corrélations Élémentaire
DIPE
Determination of Input Parameters Empirical properties
FFTBM
Fast Fourier Transformation Based Method
HTC
Heat Transfer Coefficient
IT
Integral Test
IPREM
Input Parameter Range Evaluation Methodology
IUQ
Inverted Uncertainty Quantification

ABSTRACT

PREMIUM (Post BEMUSE Reflood Models Input Uncertainty Methods) was an activity launched with the aim of pushing forward the methods of quantification of physical model uncertainties in thermal-hydraulic codes. The benchmark PREMIUM was addressed to all who apply uncertainty evaluation methods based on input uncertainties quantification and propagation. The benchmark was based on a selected case of uncertainty analysis application to the simulation of quench front propagation in an experimental test facility. Applied to an experiment, enabled evaluation and confirmation of the quantified probability distribution functions on the basis of experimental data. The scope of the benchmark comprised a review of the existing methods, selection of potentially important uncertain input parameters, quantification of the ranges and distributions of the identified parameters using experimental results of tests performed on the FEBA test facility, verification of the performed quantification on the basis of tests performed at the FEBA test facility and validation on the basis of blind calculations of the Reflood 2-D PERICLES experiment. The benchmark has shown dependency of the results on the applied methodology and a strong user effect. The conclusion was that a systematic approach for the quantification of model uncertainties is necessary.

* Corresponding author.

Email address: Tomasz.Skorek@grs.de (T. Skorek)

LB
 Large Break
 LOCA
 Loss of Coolant Accident
 MCDA
 Model Calibration through Data Assimilation
 MFBT
 Minimum Film Boiling Temperature
 PCT
 Peak Cladding Temperature
 QF
 Quench Front
 SET
 Separate Effect Test
 TH
 Thermal-Hydraulic

Keywords

Model uncertainties
 Uncertainties quantification
 Input uncertainties propagation
 Thermal-hydraulic codes
 Combined effect tests

1. Introduction

The identification of the input uncertainties and probabilistic quantification of their uncertainty are essential for the uncertainty and sensitivity analyses. The results of the BEMUSE project have shown that the input uncertainties quantification is of great importance for BEPU (Best Estimate Plus Uncertainty) analyses (CSNI, 2011). Whereas, the quantification of uncertainties related to the analysed facility and initial and boundary conditions of the transient is mainly the question of information that usually can be obtained, the quantification of physical model uncertainties is an extensive process involving experimental data, numerical methods and expert judgement. Since the model uncertainties are frequently the dominant factors in uncertainty analyses, special attention has to be paid to the methodologies applied for their quantification.

The basis for physical model uncertainties quantification is the evaluation of Separate Effect Tests (SET). Also, Combined Effect Tests (CET) and Integral Tests (IT) can be applied for the quantification. Other information sources are experience from code validation, survey of expert state of knowledge, published data about model uncertainties and if necessary, application of theoretical limitation. The preferable way of model uncertainties quantification is the comparison of code predictions with experimental data from separate effect tests. However, there are some phenomena where no single effect tests are available. An example of such a phenomenon is reflooding of the reactor after core dry out. On the basis of selected reflooding tests, a benchmark has been defined with the aim of pushing forward the quality of physical model uncertainties quantification. The main objective of the benchmark was to compare existing approaches, to estimate their quantification capabilities and to find out proper methods to quantify uncertainties of models describing phenomena for which no single effect tests exist. In the case of reflooding experiments, SETs required for validation of code models calculating enhancement of heat transfer very close to the quench front and the relative velocity downstream from the quench front do not exist. The influence of these models can only be seen at their effect on quench front propagation which is measured. Sophisticated methods are needed to determine these input uncertainties.

For performing a successful check of model uncertainty quantification, it is necessary that other potentially important uncertainties, like uncertainties of spatial modelling can be eliminated. This is the case for relatively simple test facilities, where 1-D approximation is suitable, and no particular problems should arise by discretization and development of an input data set.

In this context reflooding experiments appears to be suitable for the benchmark application:

- Some reflooding tests are available.
- Geometry of the test section is quite simple and average experienced user should not have any problem with its correct simulation.
- In reflooding only a few physical phenomena are involved, it means it should be possible to identify the reason of differences comparing the results of different calculations (participants) with experimental data.

FEBA (Ihle and Rust, 1984a, b) experiment was proposed as support for input uncertainties quantification. For the confirmation/validation phase in the proposed benchmark the PERICLES (Deruaz et al., 1985) experiment was selected. Both tests cover the similar fields of application. The PERICLES test facility consists of a test section with a larger number of fuel rod simulators. Construction of the PERICLES test facility enables investigation of radial power distribution effects. The selected test sequence follows typical way of uncertainty analysis. Input uncertainties are quantified on the basis of test facilities and applied for reactor geometry of the much larger scale. At PERICLES test facility parallel to tests with radial power distribution, tests with uniform radial power distribution were performed, where the 2D effect should be minimized. Therefore, it should be possible to separate this effect by analysis of benchmark results.

Moreover, Reflood is one of the most important thermohydraulic processes occurring in a large break loss-of-coolant (LBLOCA) scenario in PWR plants. It has significant influence on the maximum temperature in the core and therefore is of interest for safety analyses. This was recognized as an additional reason for potential participants to take part in the benchmark.

Sixteen institutions participated in the PREMIUM benchmark. One of participants used two codes and another one applied two different quantification methods. A few institutions participated only in subset of benchmark phases. In Table 1 the participating institutions, thermal-hydraulic (TH) codes and quantification methods they have used, and participation in essential phases of the benchmark are listed.

The PREMIUM benchmark was structured in five consecutive phases:

- Phase I: Final specification of the benchmark. The available methods of uncertainty quantification were described. The specification of the

Table 1
List of PREMIUM participants, used codes, and applied quantification methods.

User	Country	Code	Method	Phase		
				II	III	IV
Bel V	Belgium	CATHARE2 V25_2 mod8.1	CIRCÉ	yes	yes	yes
CEA	France	CATHARE2 V25_2 mod8.1	CIRCÉ	yes	yes	yes
CVRez	Czech Republic	RELAP5 mod3.3	CIRCÉ		yes	yes
GRS	Germany	ATHLET 2.2B	GRS, IUQ method	yes	yes	yes
IRSN	France	CATHARE2 V25_2 mod8.1	DIPE	yes	yes	yes
KAERI-1	Republic of Korea	MARS-KS1.3-COBRA-TF	CIRCÉ	yes	yes	yes
KAERI-2	Republic of Korea	MARS-KS1.3-COBRA-TF	MCDA	yes	yes	yes
KINS	Republic of Korea	MARS-KS-0003 PREMIUM version	CIRCÉ	yes	yes	yes
KIT	Germany	TRACE Version 5 patch3	IPREM	yes	yes	*
NRI	Czech Republic	ATHLET 2.1A	-	yes		
OKBM-1	Russian Federation	KORSAR/BR	CIRCÉ		yes	yes
OKBM-2	Russian Federation	RELAP/SCDAPSIM/mod3.4	CIRCÉ	yes	yes	yes
PSI	Switzerland	TRACE V5.0P3-UQ	Expert judgement based method			yes
SJTU	China	RELAP5/SCADPSIM/mod3.4	IPREM		yes	yes
Tractebel	Belgium	RELAP5 mod3.3	IUQ method	yes	yes	yes
UNIFI	Italy	RELAP5 mod3.3 patch 3	IPREM	yes	yes	yes
UPC & CSN	Spain	RELAP5 mod3.3 patch 4	CIRCÉ	yes	yes	yes
VTT	Finland	APROS 5.11.2	CIRCÉ (bias) + FFTBM (range)	yes	yes	yes

FEBA test facility and experimental data from the Series I were distributed among participants.

- Phase II: Identification and initial quantification of important input uncertainties for reflood prediction. This was performed on the basis of the results of Test 216 of series I in reflood FEBA experiments.
- Phase III: Quantification of physical model uncertainties identified in Phase II, using measured data from 6 FEBA tests of the series I (including Test 216) and selected quantification method.
- Phase IV: Verification and validation of the results of Phase III, by propagation of the input uncertainties through simulation model (uncertainty analyses of reflood test simulations) and comparison to the measured data of the chosen reflood experiments. At first, the consistency of quantified uncertainty was verified by uncertainty analyses of the 6 FEBA tests used in the quantification. Secondly, the same input uncertainties were propagated through the simulation models by calculation of another reflood experiment, PERICLES 2D. These uncertainty analyses were performed blindly.
- Phase V: Compilation and comprehensive analysis of the results. Lessons learned were identified and conclusions were drawn. The report on the PREMIUM benchmark was the final product of this phase.

2. Experiments chosen for the PREMIUM benchmark

For the PREMIUM benchmark two experiments were selected. For the quantification step. The tests from the FEBA program have been chosen. However, if any participants prefer, they could use their own reflood experiment, provided that it was sufficiently validated, and they accept to make their experimental results available for other participants. The same experiment has been applied for the verification step.

For the validation step six test from the 2-D PERICLES reflood experiment have been selected. The selected test runs have not been published at that time. This enabled performing of blind analyses as it was desired in the benchmark specification.

2.1. The FEBA experimental program

The FEBA experiment is devoted to the study of the reflood. The experimental data of the six unblocked FEBA tests of the Series I were provided to the participants, in order to derive the uncertainty of the physical models influential during reflood. The test section consisted of a full-length 5×5 rod bundle of electrically heated rods with PWR fuel rod dimensions, surrounded by a housing insulated to reduce heat losses to environment, as shown in Fig. 1.

The power profile was of cosine type approximated by 7 steps of different power density in axial direction. Seven regularly located spacer grids were placed in the bundle.

Prior to the test run, the fuel rod simulators were heated in stagnant steam to desired initial cladding temperature of roughly 800°C , using a low rod power. In the meantime, the test bundle housing was heated passively to the targeted initial temperature of roughly 635°C by radiation from the rods. The aim of choosing a thick wall was to prevent premature quenching of the wall relative to the bundle quench front progression.

At the beginning of the test run, the bundle power was increased to a level corresponding to a 120% ANS-standard at about 40 s after reac-

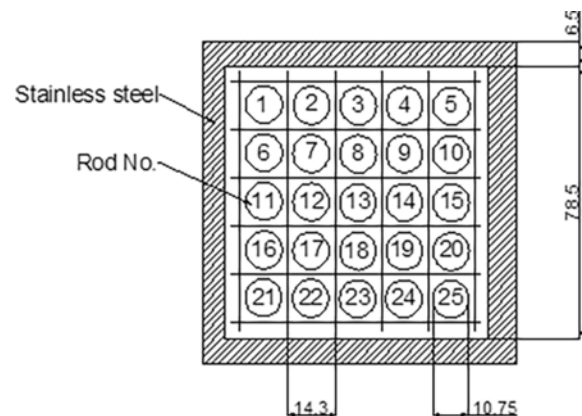


Fig. 1. FEBA rod bundle. Cross-section view.

tor shut down. Simultaneously, the cold water supply was activated (its temperature is about 40 °C).

The inlet velocity and the pressure were varied for the 6 considered tests according to Table 2.

2.2. The PERICLES experimental program

PERICLES was designed to investigate 2-D effects which can occur in a PWR core where the rod power is not identical from one assembly to the other ones. The experiment consists of three different assemblies, denoted here by A, B and C (Fig. 2). These assemblies are contained in a vertical cold housing with a rectangular section. Each assembly contains $7 \times 17 = 119$ full length heater rods, so that the total number of heater rods is 357.

The rods were heated by two independent electrical power sources, creating the possibility to heat the central assembly (the ‘hot’ assem-

Table 2
Boundary conditions of the 6 FEBA tests considered for PREMIUM Phase III.

Test No.	Inlet water velocity (cm/s)	System pressure (bars)
223	3.8	2.2
216	3.8	4.1
220	3.8	6.2
218	5.8	2.2
214	5.8	4.1
222	5.8	6.2

bly) more than the two lateral ones: A and C (the ‘cold’ assemblies). The axial power profile is, as for FEBA, of cosine type, with 11 levels.

The experimental procedure is slightly different from that of FEBA, since, at the beginning of the transient, the whole power was immediately switched on until the target initial maximum value of the clad temperature in the hot assembly was reached (generally 600 °C). The outer part of the housing was heated, with the aim of maintaining its temperature a few degrees above the saturation temperature.

The six tests considered for PREMIUM benchmark are listed in Table 3. The RE0064 test is the reference test. In other tests, only one boundary condition was modified with respect to RE0064. In Table 3, $HF_{nom}(HA)$ and $HF_{nom}(CA)$ are the nominal heat fluxes in the Hot and Cold Assemblies respectively, F_{xy} is the radial power peaking factor, i.e. the ratio $HF_{nom}(HA)/HF_{nom}(CA)$, GO is the inlet mass velocity entering the bottom of each assembly during the reflood stage, T_{wi} is the initial cladding temperature in the middle of each assembly (at the beginning of the reflood stage). The value for the hot assembly must be reached in the calculation to launch the reflood. The value of T_{wi} in the cold assemblies should also be reached if possible. DT is the subcooling of water entering the assemblies.

3. Phase I: Introduction and methodology review

The main objective of Phase I was detailed specification of the benchmark and review of the methods which could be applied by participants for uncertainties quantification. The simultaneous quantification of the several model uncertainties on the basis of combined effect test is a difficult procedure and only few statistical methods were available at that time. They were:

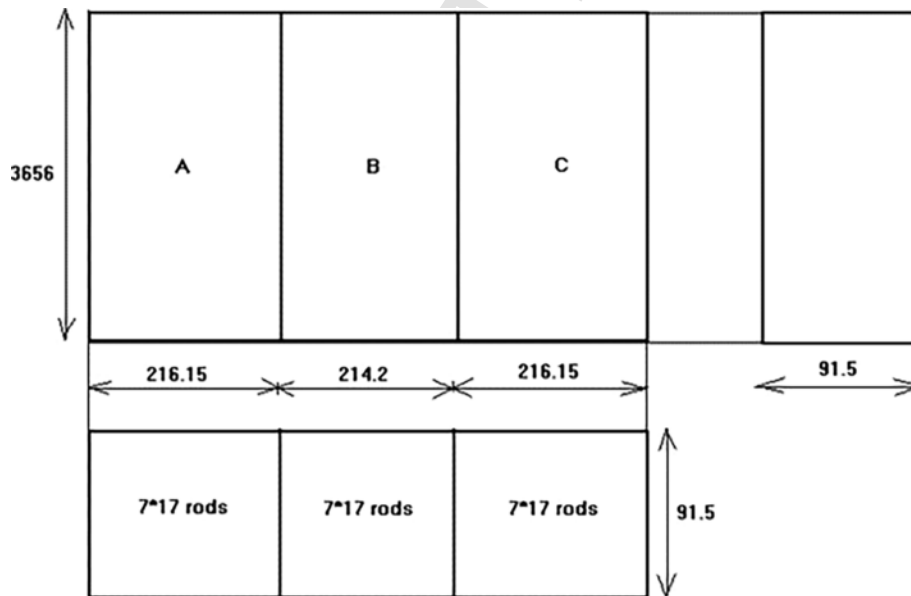


Fig. 2. The 2-D PERICLES experiment (dimensions indicated in mm).

Table 3
Experimental conditions (HA: hot assembly; CA: cold assembly).

Test No	$HF_{nom}(HA)$ W/cm ²	$HF_{nom}(CA)$ W/cm ²	F_{xy}	$GO(HA)$ g/cm ² s	$GO(CA)$ g/cm ² s	$T_{wi}(HA)$ °C	$T_{wi}(CA)$ °C	DT °C	P (bar)
RE0062	2.93	2.93	1	3.6	3.6	600	600	60	3
RE0064	4.2	2.93	1.435	3.6	3.6	600	475	60	3
RE0069	2.93	2.93	1	3.6	3.6	475	475	60	3
RE0079	4.2	2.93	1.435	3.6	3.6	600	475	90	3
RE0080	4.2	2.93	1.435	5	5	600	475	60	3
RE0086	4.2	2.93	1.435	3.6	3.6	600	475	60	4

- CIRCE method (Calcul des Incertitudes Relatives aux Corrélations Élémentaires) based on the maximum likelihood which quantifies the parameters under the assumption of their linear dependency on the model outputs, developed at CEA (de Crécy, 2001),
- Method based on data assimilation and covariance matrixes developed at KIT (Cacuci and Ionsecu, 2010),
- IPREM (Input Parameter Range Evaluation Methodology) based on FFTBM (Fast Fourier Transformation Based Methods) shortly before developed at University of Pisa (UNIPI) (Kovtonyuk et al., 2017)

Other existing possibilities were approaches based on extensive using of expert judgement like, “trial-and-error” method by performing calculations of selected experiments related to the investigated phenomenon, definition of “biases”, evaluation of information and experience from model development.

Developers of the CIRCE and IPREM methods CEA and University of Pisa were participants of the benchmark and offered the related software to the participants of the benchmark, who wanted to apply their methods for quantification of model uncertainties. The development of the covariances matrices is a very complicated and time consuming process, so its application was not taken into account by any of benchmark participants. The participant representing KIT decided to use IPREM method.

The participants can be divided into four groups according to the uncertainties quantification:

- Participants having at their disposal methodologies for quantification of uncertainties of the physical models: CEA, UNIPI
- Participants who became users of the available methodologies: Bel V, SJTU, CVRez, VTT, KIT, KINS, OKBM and UPC.
- Participants who developed their own method in parallel with PREMIUM participation: VTT, Tractebel, IRSN, GRS, KAERI
- Participants who used an expert-judgment based method improved with methods of fitting of data: PSI

In the course of the benchmark the participants who declared development of their own methods undertaken this effort and developed new methodologies. KAERI developed and applied MCDA (Model Calibration through Data Assimilation method (Heo et al., 2014). IRSN developed DIPE (Determination of Input Parameters Empirical properties) method (Joucla and Probst, 2008). TRACTEBEL has developed a sampling based inverse uncertainty quantification (IUQ) method using of DAKOTA software (Zhang et al., 2019). GRS developed a method based on combination of separate test evaluation and inverted uncertainty method for combined effect test evaluation (Skorek, 2017). VTT developed an approach where the combination of available methods was used: CIRCE for calibration (definition of CIRCE calibration in §6.1.1) and IPREM for uncertainty range determination.

Concerning the experimental basis of the benchmark it was agreed that all participants would only use the proposed experiments (FEBA and PERICLES). None of the participants decided to use, in addition, other reflood experiment for the quantification of the uncertainties.

The list of 16 institutions that decided to participate in the benchmark, grouped in 15 work teams (CSN and UPC participate as a single team), together with the system thermohydraulic code and the quantification method used by each participant can be seen in Table 1.

4. Phase II: Identification of influential input parameters.

The main goal of Phase II was the identification of input parameters (IP) important for reflood simulation. The influential IP identification

has been performed based on the Test 216 of FEBA experimental program.

The first step was development of input data and performing of the reference calculation for the selected test 216. Thirteen organizations were involved in Phase II, using 8 different codes (see Table 1). All of them were 1D system TH codes, except one, COBRA-TF which is a sub-channel module of the system code MARS-KS.

Most of participants modelled the test section of FEBA as a single vertical channel and a single heat rod/heat structure. The sole user of TRACE was KIT, and applied a CHAN component, representing a 5×5 bundle. KAERI modelled 1/8 of the bundle with COBRA-TF. All participants included a model of the test section housing. Regarding the spacer grids, some participants represented them by a flow area reduction and activated special models for heat transfer enhancement; others simply applied form loss coefficients.

The number of axial nodes used by the participants ranged from 20 to 78. This number does not take into account the possible refinement performed by the codes in the vicinity of the quench front, as a part of their reflood models.

The results for quench front propagation and peak cladding temperature are shown in Figs. 3 and 4, respectively.

The base case calculations show spread in predicted cladding temperature and quench front propagation, with respect to measured data. All participants except one overpredicted QF progression. There are some discrepancies in the modelled initial conditions. Some participants simulated the heat-up phase, while others initialized the models at the conditions reported for the beginning of the transient. The time trends calculated for cladding temperatures by most of participants show oscillatory behaviour (probably having a numerical origin).

Most of participants obtained satisfactory values for peak cladding temperature and bundle rewet time. RELAP and CATHARE users generally underpredicted the PCT. The APROS simulation show strong underprediction of the PCT, while the TRACE user obtained the maximum overprediction. The predicted bundle quench times show a significant spread up to 30% with respect to the measured value. However, some of the calculated results showed considerable discrepancy regarding the measured data, they were considered as a suitable basis for performing the sensitivity analysis.

The next step was the selection of the most influential input parameters by performing the sensitivity studies. This step consisted of identification, by each participant, of the physical models included in their codes which are influential in the reflooding scenario, preliminary quantification of their uncertainty (in terms of range of variation), and selection the important uncertain input parameters.

The coordinator of the Phase II, University of Pisa proposed a procedure for identification and selection of influential uncertain input parameters (Kovtonyuk et al., 2015). Nevertheless, the majority of participants applied their own procedures. However, all applied procedures considered sensitivity analysis of the reference case Test 216 of FEBA.

From the total number of input uncertainties identified by each participant as possibly influential, according to results of sensitivity studies the most important input parameters have been selected. The majority of participants considered about 20 parameters, except VTT and KIT, who initially considered 40 and 56 parameters respectively.

The results of sensitivity analyses depend mainly on the physical model, which are usually different in different codes, and on the range of variation, i.e. preliminary quantification of the investigated parameters. Whereas, the dependency on the model remains constant for the code, the range of variation is strongly user dependent. As it can be seen on example of interfacial friction (see Fig. 5), even for the same code the range of the variation can differ between the various user up to the order of magnitude. This can strongly influence the results of the selection of the most influential parameters.

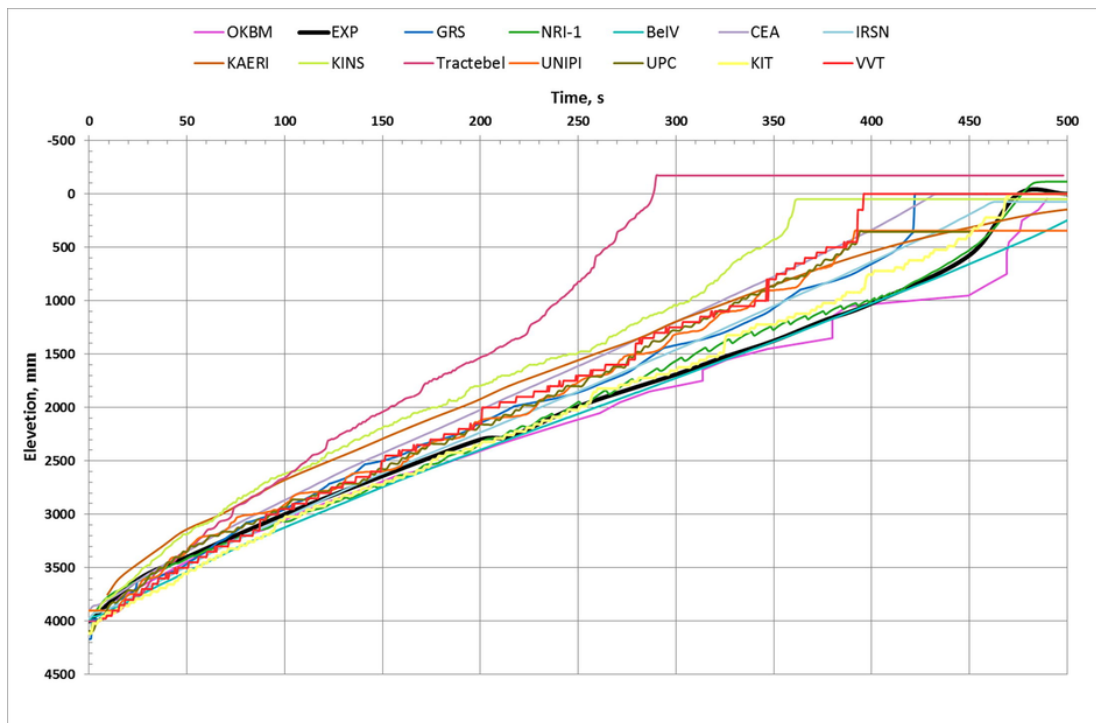


Fig. 3. FEBA: Quench front propagation in Base Case.

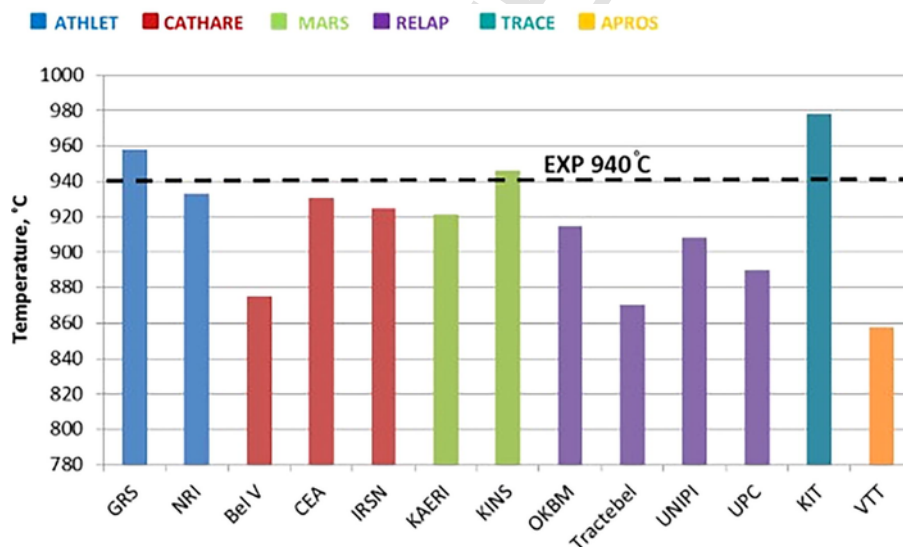


Fig. 4. PCT comparison in Base Case.

Out of the total set of 72 influential parameter, initially considered by all the participants, only 6 were identified as influential by at least 5 participants:

- Bundle power (density)
- Wall heat transfer coefficient (HTC): some participants separated the HTC for liquid and for vapor. Some codes showed very low sensitivity of the time of rewet with respect to this parameter.
- Interfacial friction coefficient: some participants distinguished between the friction coefficient for bubbles and droplets, and for dispersed vapor. This parameter influences the cladding temperature in an indirect way, through the void fraction in front of the fuel rods.
- Interfacial HTC.
- Droplet diameter.

- Heat transfer enhancement at the quench front: the actual parameters identified are code specific and may have different influence on calculation results.

The results of the performed parameter selection are shown in Fig. 6 in terms of the reduction of the considered parameter numbers.

5. Phase III: Quantification of model uncertainties on the basis of FEBA reflood experiments.

The goal of Phase III is the quantification of model uncertainties on the basis of CETs. Such procedure is addressed in particular to uncertainties quantification of models related to phenomena for which SETs do not exist, as it is the case for heat transfer enhancement at the quench front during the reflood phase of a Large Break LOCA.

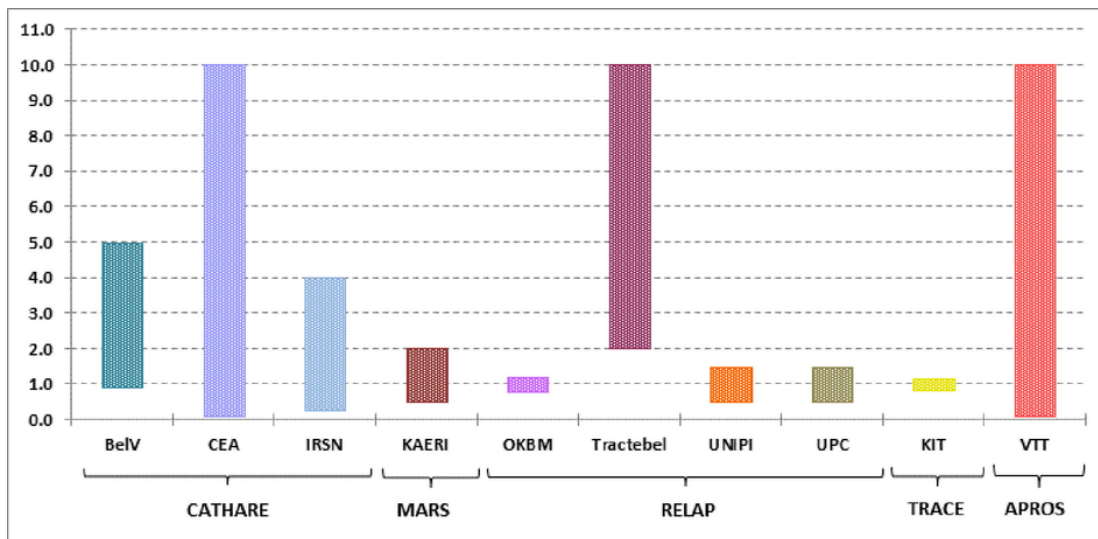


Fig. 5. Interfacial friction coefficient variation range (multiplier).

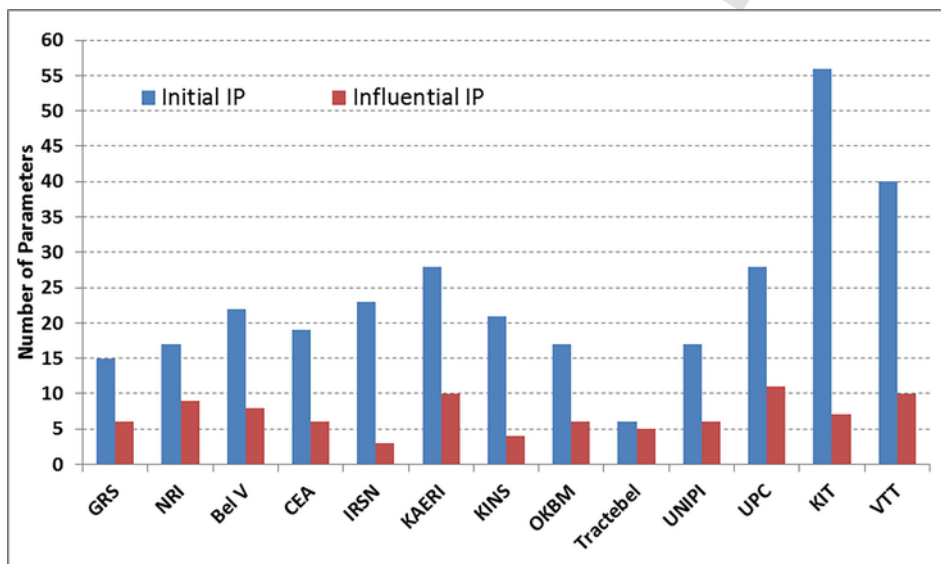


Fig. 6. Selection of influential IP by participants.

The coordinator of Phase III was GRS (Germany). Altogether 14 organizations have participated in Phase III supplying 16 contributions. The participants used different thermal-hydraulic system codes. As it can be seen in Table of Participants in PREMIUM benchmark two codes have been used by several participants. RELAP5 MOD3.3 has been used in 5 contributions, and CATHARE V2.5 has been used by 3 participants.

The information necessary for PREMIUM was extracted from the KfK experimental reports (Ihle and Rust, 1984) by the Phase III coordinator, summarized and supplied to the participants. The measured data were available only in the form of plotted curves, so an important part of the task was the digitalization of such curves.

As it can be seen in Table 4, the majority of participants applied one of the methods offered in the frame of PREMIUM. Eight participants applied CIRCÉ method and three participants applied FFTBM. One participant (VTT Finland) applied a combination of these two methods. CIRCÉ method has been applied by VTT for determination of biases and FFTBM for the determination of the uncertainty ranges. The remaining four participants applied each one their own method.

Two organizations (KAERI and OKBM) submitted two contributions each. OKBM used the quantification method CIRCÉ with two different

codes RELAP and KORSAR. KAERI used for code COBRA two different quantification methods CIRCÉ and MCDA

Phase III has been divided into three steps:

Step 1:

Definition of the list of parameters to be considered, and initial quantification.

The basis for the definition of parameters to be considered in the Phase III evaluation/quantification step were the results obtained by each participant within the Phase II. The results of the Phase II enabled comparison with results obtained by other participants, in particular those using the same thermal-hydraulic code. This gave an opportunity for critical review of their own list of selected uncertainties and preliminary quantification ranges on the basis of which the selected parameters had been found as influential. A critical review of the results obtained within the Phase II should lead to improvement of the list of selected parameters which were going to be quantified within the Phase III by each participant

Step 2:

Uncertainty quantification for the selected model parameters. Each participant was responsible for choosing an adequate method of model

Table 4
Methods and chosen experimental tests & responses in Phase III.

Particip.	Code	Method	Tests used	Responses used
CVRez	RELAP	CIRCÉ	FEBA tests: 223, 216, 220, 218, 222	clad temp., quench time, Δp
UPC	RELAP	CIRCÉ	all 6 FEBA tests	clad temp., water carried over, quench time
OKBM-1	RELAP	CIRCÉ	FEBA tests: 223, 216, 220, 218, 214	clad temp., quench time
OKBM-2	KORSAR	CIRCÉ	all 6 FEBA tests	clad temp., quench time
CEA	CATHARE	CIRCÉ	all 6 FEBA tests	clad temp., quench front elevation
Bel V	CATHARE	CIRCÉ	all 6 FEBA tests	clad temp., quench time
KINS	MARS-KS	CIRCÉ	all 6 FEBA tests	clad temp.
KAERI-1	COBRA	CIRCÉ	FEBA tests: 214, 216, 218, 223	clad temp.
KAERI-2	COBRA	MCDA	FEBA tests: 214, 216, 218, 223	clad temp.
VTT	APROS	FFTM/CIRCÉ	all 6 FEBA tests	clad temp., housing temp., quench time
UniPisa	RELAP	FFTBM	FEBA test 216	clad temp., quench time
SJTU	RELAP	FFTBM	FEBA test 216	clad temp., quench time, Δp
KIT	TRACE	FFTBM	FEBA test 216	clad temp., quench time
IRSN	CATHARE	DIPE	all 6 FEBA tests	clad temp., quench front propagation
GRS	ATHLET	IUQ.	all 6 FEBA tests	Δp , water carried over

Table 4 (Continued)

Particip.	Code	Method	Tests used	Responses used
			FEBA tests: 216, 223	clad temp.
Tractebel	RELAP	IUQ	Separate effect tests all 6 FEBA tests	clad temp.

uncertainties quantification. The results of the Phase II sensitivity analyses and experience concerning test run simulation indicate that the experiment and test facility related uncertain input parameters are secondary compared to the physical model parameters. Nevertheless, consideration of influence of uncertain parameters different to model parameters in the quantification process was left to individual participant decision.

Step 3:

Preliminary check of quantified uncertainty ranges. Uncertainty and sensitivity analysis of the Test run 216 from the Series I of FEBA had to be performed. It is the FEBA test most similar to the PERICLES experiment, from the standpoint of boundary conditions. The selection of various calculated quantities for comparison with measured data, and not only cladding temperatures, should give an opportunity to find out if compensating errors take place by the test run simulation. Participants could improve their uncertainty quantification, if it was found necessary.

5.1. Experimental data used for uncertainties quantification

For performing a successful check of model uncertainty quantification method, it is of importance that other potentially important uncertainties, like uncertainties of spatial modelling can be eliminated. This is the case for relatively simple test facilities, where 1-D approximation is adequate, and no particular problems should arise by discretization and development of an input data set. In this context FEBA reflooding experiments appears to be suitable for the benchmark application.

For the quantification of reflood model uncertainties, all participants have used the tests from Series I of FEBA experiments (Table 4). Participants were allowed using their own reflood experiments (provided that they are sufficiently validated and made available to other participants). The measured data of FEBA Series II could have been additionally used in the task of quantification. Despite this flexibility neither other reflooding experiment nor tests from Series II were used by the participants in this Phase.

Majority of participants used the 6 tests of Series I. Only 3 used solely test 216. They were FFTBM users, and the reason was that the supplied software enabled the application of the method to one test run, only. VTT made an improvement to the software to apply it to several tests. Thus, VTT could use the 6 tests of the Series.

Regarding the type of measured responses used for the quantification, all participants considered cladding temperatures, and almost all of them used quench front progression (quench time or QF elevation), too. Some participants considered also pressure drop measurements, and only few used the measured data of water carried over. Only one (VTT) included the measured housing temperature.

Only one participant (GRS) made use of data from SETs, in order to quantify a part of the model uncertainties (two correlations related to wall heat transfer at dry-out condition). The goal of this procedure was

to benefit from the large database for wall heat transfer correlations to accomplish a better quantification of the uncertainty.

5.2. Input parameters quantified within Phase III

The input parameters that have been quantified during Phase III are listed in Table 5. The number of parameters quantified by each participant varied between 2 and 6. They were all model parameters, with an exception. One participant (KIT), in addition to 6 model parameters, made the quantification of uncertainty of the rod bundle power, which in fact is a boundary condition of the experiment, having its own experimental uncertainty.

The number of parameters quantified by RELAP users varies from 2 to 5. On the other hand, all users of CATHARE codes considered three parameters. Regarding the quantification method the majority of CIRCÉ users considered two or three parameters. Only KAERI considered 4 parameters. Users of other methods considered generally larger number of parameters.

All participants considered uncertainties related to the wall heat transfer. As it could be expected, all participants considered uncertainty of heat transfer correlation at dry-out conditions. Almost all participants considered also uncertainties of momentum equation closure relations. The interfacial heat transfer uncertainty has been considered by 11 participants. Usually, both the uncertainty of interfacial momentum transfer and interfacial heat transfer were considered. Some participants considered only uncertainty of momentum equation closure relation. Only two participants considered uncertainty of interfacial heat transfer not taking into account interfacial momentum transfer uncertainty.

The following heat transfer related parameters were considered by participants:

- Film boiling heat transfer coefficient (HTC) – 10 participants
 - Film boiling HTC total – 2 participants
 - Film boiling HTC gas/liquid phase separately – 8 participants
- General HTC for dry-out condition (above the quench front) – 2 participants
- Global HTC for all heat transfer regimes (applied in the whole bundle) – 2 participants
- HTC for steam convection – 3 participants
- Minimum film boiling temperature – 2 participants
- Heat transfer enhancement at quench front – 5 participants

- Heat transfer enhancement at grid spacers – 3 participants

The interfacial heat transfer parameters considered by participants are following:

- Global interphase heat transfer – 4 participants
- Interphase heat transfer for mist flow – 6 participants
- Droplet diameter (key parameter used in evaporation model) – 3 RELAP users

The considered momentum equation closure relation uncertainties are:

- Global interfacial friction – 8 participants
- Interfacial friction for mist flow – 5 participants
- Droplet diameter (key parameter used in interfacial drag model) – 2 RELAP users
- Entrained liquid fraction – 1 participant
- Wall friction of liquid phase – 1 participant

Although the variety of the considered parameters is limited, even the users of the same code considered frequently different parameters. The code RELAP was used by 6 participants. Only UNIPI and SJTU considered the same 4 parameters out of total number of 5 and 4 parameters considered by them respectively. UNIPI and TRACTEBEL considered the same 3 parameters out of total number 5 and 4 parameters. In the case of CATHARE CEA and IRSN considered the same 3 parameters. The third user of CATHARE (Bel V) considered only 1 parameter in common with CEA and IRSN. On the other hand, GRS and VTT using different codes considered three common parameters (parameters related to the same or equivalent models) out of total of six parameters used by each of them.

The type of probability density functions (pdf's) applied by the participants is clearly method dependent. As it can be seen in Table 5 all the users of the CIRCÉ method and similar data assimilation method applied normal or log-normal probability distribution functions. The IRSN applied histograms. All other participants applied uniform distributions. The reasons for the selection of the type of distribution are the assumptions making the basis of the methods. In CIRCÉ formulation normal or log-normal distribution probability distribution functions of the quantified parameters have been assumed. In other methods, the uniform distribution was applied. Even in the IRSN method, where an

Table 5
Uncertain input parameters considered by participants and their probability distributions.

Particip.	Code	No of par.	Wall HTC	Interfacial HTC	Momentum eqn – closure rel.	Method
CVRez	RELAP	2	1 [log-norm]	–	1 [log-norm]	CIRCÉ
UPC	RELAP	3	1 [log-norm]	1 [log-norm]	1 [log-norm]	CIRCÉ
OKBM-1	RELAP	3	2 [log-norm]	1 (droplet) [log-norm]	–	CIRCÉ
OKBM-2	KORSAR	2	1 [log-norm]	–	1 [log-norm]	CIRCÉ
CEA	CATHARE	2 + 1(bias)	1 [log-norm] + 1(quen.)	–	1(mist flow) [log-norm]	CIRCÉ
Bel V	CATHARE	3	1 + 1(quen.) [log-norm]	–	1 [log-norm]	CIRCÉ
KINS	MARS-KS	2	1 [norm]	1 [norm]	–	CIRCÉ
KAERI-1	COBRA	4	1 + 1(grid) [log-norm]	1(mist flow) [log-norm]	1(mist flow) [log-norm]	CIRCÉ
KAERI-2	COBRA	4	1 + 1(grid) [log-norm]	1(mist flow) [log-norm]	1(mist flow) [log-norm]	MCDA
VTT	APROS	6	2 + 1 [norm, log-norm log-norm]	1 [log-norm]	1 [log-norm] + 1 (wall frict.) [norm]	FFTBM/CIRCÉ
UniPisa	RELAP	5	2 [uniform]	1 + 1(droplet) [uniform]	1 + 1(droplet) [uniform]	FFTBM
SJTU	RELAP	4	2 [uniform]	1(droplet) [uniform]	1 + 1(droplet) [uniform]	FFTBM
KIT	TRACE	6 + rod power	3 + 1(grid) [uniform]	1 [uniform]	1 [uniform]	FFTBM
IRSN	CATHARE	3	1 + 1(quen.) [histogram]	–	1(mist flow) [histogram]	DIPE
GRS	ATHLET	6	2 + 1(quen.) [uniform]	1 [uniform]	1 + 1 (entr.) [uniform]	IUQ
Tractebel	RELAP	4	1 [uniform]	2 [uniform]	1 [uniform]	IUQ

empirical distribution could be obtained in the course of quantification the histogram distribution has been assumed.

The normal and log-normal distributions determined by CIRCE are not truncated. The truncation of distributions was left to the users' decision. Some users of CIRCE performed this step and truncated the obtained distributions, while others applied the non-truncated distributions. The truncation was performed at 2.5% and 97.5% percentiles of the distribution. The truncation values of the distributions are presented in the tables as "Min" and "Max" values of the variation range. Also, those users of CIRCE who did not perform the truncation supplied for comparison the values of 2.5% and 97.5% percentiles of their distribution (also presented in Tables 6 and 8).

Application of non-truncated distributions can lead to generation of extreme values of parameters by sampling. Parameters combinations with such extreme values are extremely improbable in random sampling, but they could result in failure of the code run or production of non-physical code results. Application of truncated distributions prevents generally generation of such extreme elements of the sample. Also, by application of truncated distributions there are still differences between normal and uniform distributions. In random sample generation for normal distributions, elements near the mean value of the distribution are preferred. For uniform distributions, the probability of el-

ement generation is equal for the whole range of variation. Using of higher order of Wilks' formula (Wilks, 1941) can reduce to some extent the differences due to application of different probability distribution functions, particularly those due to application of non-truncated distributions.

5.3. Consideration of uncertainties other than physical model uncertainties

PREMIUM is focused on the estimation of uncertainties of model parameters. The uncertainty of other type of input parameters (initial and boundary conditions, material properties...) should be estimated from other sources (e.g. measurement devices). In the specification of Phase III information about experimental uncertainties and thermal properties uncertainties in FEBA experiment were supplied.

The information in Phase III specification concerning the thermal properties of materials used in FEBA test facility was not obtained from experimenters but was the result of Phase III coordinator survey of literature and estimations. Therefore, they could be subject to corrections and modifications by the participants. However, taking into account experience from parametric sensitivity study performed within the Phase II, it seems that in the case of FEBA experiment the uncertainties of ma-

Table 6
Uncertainty ranges applied for heat transfer at dry out conditions.

Particip.	Code	Method	Wall HTC	Min	Max	Ref./calibr value
CVRez	RELAP	CIRCE	1 (all regimes)	0.433	0.954	0.643 (calibr.)
UPC	RELAP	CIRCE	1 (film-liquid)	0.75	1.37	1.01 (calibr.)
OKBM-1	RELAP	CIRCE	2 (film-gas/liq)	0.43/1.06	0.97 /1.49	0.646/1.272 (calibr.)
OKBM-2	KORSAR	CIRCE	1 (film-gas)	0.816	1.34	1.046 (calibr.)
CEA	CATHARE	CIRCE	1 (dry out)	0.73	1.44	1.03 (calibr.)
Bel V	CATHARE	CIRCE	1 (all regimes)	0.692	1.780	1.110 (calibr.)
KINS	MARS-KS	CIRCE	1 (MFBT)	(0.493)	(0.891)	0.692 (calibr.)
KAERI-1	COBRA	CIRCE	1 (film - liq.)	0.52/0.90	1.02 /1.18	0.73/1.03 (calibr.)
KAERI-2	COBRA	MCDA	1 (film - liq.)	0.60/0.80	1.26 /1.26	0.86/1.01 (calibr.)
VTT	APROS	FFTM/CIRCE	2 (steam + MFBT)	0.05 /(0.65)	1.7/(1.7)	0.7/1.2 (calibr.)
UniPisa	RELAP	FFTBM	2 (film-gas/liq)	0.35/0.50	2.8/1.3	1.0/1.0
SJTU	RELAP	FFTBM	2 (film-gas/liq)	0.15/0.77	1.92 /1.44	1.0/1.0
KIT	TRACE	FFTBM	2 (steam + film)	0.41/0.36	1.4/1.4	1.0/1.0
IRSN	CATHARE	DIPE	1 (dry out)	0.58	1.56	1.0
GRS	ATHLET	IUQ	2 (steam + film)	0.85/0.65	1.25/1.3	1.0/1.0
Tractebel	RELAP	IUQ	1 (film-gas)	0.7	1.3	1.0

Table 7
Uncertainty ranges applied for interfacial heat transfer model.

Particip.	Code	Method	IHT correl.	Min	Max	Ref./calibr value
CVRez	RELAP	CIRCE	-			
UPC	RELAP	CIRCE	1 (global)	0.29	2.07	0.77 (calibr.)
OKBM-1	RELAP	CIRCE	1(droplet)	0.617	0.844	0.722 (calibr.)
OKBM-2	KORSAR	CIRCE	-			
CEA	CATHARE	CIRCE	-			
Bel V	CATHARE	CIRCE	-			
KINS	MARS-KS	CIRCE	1 (mist flow)	0.148	1.967	1.058 (calibr.)
KAERI-1	COBRA	CIRCE	1 (mist flow)	0.59	1.97	1.07 (calibr.)
KAERI-2	COBRA	MCDA	1 (mist flow)	0.77	1.68	1.13 (calibr.)
VTT	APROS	FFTBM + CIRCE	1 (global)	0.05	3.5	0.5 (calibr.)
UniPisa	RELAP	FFTBM	1(global) + 1(droplet)	0.2 (0.7 mm)	5.0 (2.5 mm)	1.0 (1.5 mm)
SJTU	RELAP	FFTBM	1(droplet)	0.90 mm	2.35 mm	1.5 mm
KIT	TRACE	FFTBM	1 (global)	0.572/0.000	1.546/2.652	1.0
IRSN	CATHARE	DIPE	-			
GRS	ATHLET	IUQ	1 (mist flow)	(1.0)10 ⁹	(~3.2)10 ¹⁰	(1.0)10 ⁹
Tractebel	RELAP	IUQ	1(global) + (mist flow)	0.4 (0.4)	1.1 (1.1)	1.0 (1.0)

Table 8
Uncertainty ranges applied for momentum equation constitutive equations.

Particip.	Code	Method	Interfacial friction	Min	Max	Ref./calibr. value
CVRez	RELAP	CIRCÉ	1 (below quench)	0.487	0.906	0.664 (calibr.)
UPC	RELAP	CIRCÉ	1 (global)	0.87	1.37	1.09 (calibr.)
OKBM-1	RELAP	CIRCÉ	–			
OKBM-2	KORSAR	CIRCÉ	1 (global)	0.423	2.08	0.938 (calibr.)
CEA	CATHARE	CIRCÉ	1 (mist flow)	0.73	0.77	0.75 (calibr.)
Bel V	CATHARE	CIRCÉ	1 (global)	0.076	0.079	0.077 (calibr.)
KINS	MARS-KS	CIRCÉ	–			
KAERI-1	COBRA	CIRCÉ	1 (mist flow)	0.82	0.85	0.83 (calibr.)
KAERI-2	COBRA	MCDA	1 (mist flow)	0.49	1.49	0.84 (calibr.)
VTT	APROS	FFTBM /CIRCÉ	1(mist flow) + 1(wall friction)	1.4	6.0	2.0 (calibr.)
UniPisa	RELAP	FFTBM	1 (global) + 1(droplet)	0.86 (0.5 mm)	1.6 (1.9 mm)	1.0 (1.5 mm)
SJTU	RELAP	FFTBM	1 (global) + 1(droplet)	0.51 (0.90 mm)	1.53 (2.35 mm)	1.0 (1.5 mm)
KIT	TRACE	FFTBM	1 (global)	0.843	1.261	1.0
IRSN	CATHARE	DIPE	1 (mist flow)	0.1	7.3	1.0
GRS	ATHLET	IUQ	1 (global) + 1(entrain.)	0.64	1.60	1.0 (rel. velocity)
Tractebel	RELAP	IUQ	1 (global)	0.7	3.4	1.0

material data are considerably less influential than the physical model uncertainties.

Unfortunately, in the description of the FEBA experiment (Ihle and Rust, 1984) there is only very little information about uncertainties of the experimental data. The accuracy of the cladding and housing temperatures can be expected to be high. A typical accuracy of chromel-alumel thermocouple is about $\pm (0.4\% - 0.5\%) \cdot \text{Temp. } [^{\circ}\text{C}]$. For the measured temperature range, this corresponds to about $\pm 5^{\circ}\text{C}$. In addition, it has to be taken into account that the thermocouples measure not exactly the cladding surface temperature. The accuracy of the pressure drop measurement was not reported. However, a typical error of pressure transducers is about 1% of measured pressure range by constant temperature. Since the temperature in the FEBA experiments varied strongly along the test section the error can be much higher. The accuracy of the pressure drop measurements in other test facilities with similar bundle configuration was estimated at $\pm 10\%$. It could be a reasonable estimation also for the FEBA pressure drop measurements.

The uncertainty of the measured mass of water carried over the test section is difficult to estimate. The amount of the water carried over was measured using water collection tank. The mass of water in the tank can be measured quite precisely with estimated accuracy about 1% – 2%. However, the mass collected in the tank may be different compared to the water carried over, particularly shortly after initiation of water carryover phenomenon. At the beginning of the transient, water carried over the test section can evaporate on the hot surfaces of the upper plenum and some amount can be also accumulated in the upper plenum before water reaches the water collecting tank. A small part of water carried over the test section can be entrained by steam leaving the upper plenum. All these effects lead to underestimation of the measured mass of water carried over. The only possibility for overestimation of the water mass results from inaccuracy of water inventory measurement in the collecting tank. But this seems to be rather small. A reasonable estimation of the measured water mass uncertainty could be the range: ($\sim +0$ kg; -0.5 kg). The measured data of water carried over are available only for the initial part of the transients. The size of the tank was limited to 10 kg, and after filling of the tank, further measurements of the mass water carried over were not possible.

In addition to the uncertainties associated to the measured parameters of the reflooding in the test section, the uncertainty of the test boundary and initial conditions could be of importance for quantification of model uncertainties. The inlet velocity, inlet temperature, system pressure and bundle power are constants or evolve slowly, and thus the measurements should be quite accurate. The measurement er-

ror seems to be comparable with the error associated to numerical values obtained by digitalization of plotted curve.

The digitalisation error of the bundle power curve is about ± 1.5 kW – ± 2.0 kW. It is about 1.5–2.0% of power, the accuracy in the range of electrical power measurement error.

The inlet water temperature digitalisation error is about $\pm 2^{\circ}\text{C}$. It is also in the range of water temperature measurement accuracy. However, during initial period of the test runs, cold water filled inlet plenum, where walls had much higher temperature. This could result in non-homogeneous temperature distribution in the inlet plenum and significantly lower accuracy of the estimated inlet temperature. This is a short-term phenomenon. After 10–30 s the inlet temperature stabilised and the inlet temperature measurements are expected to reach $\pm 2^{\circ}\text{C}$ accuracy.

The system pressure measurements show pressure variation around the value defined as test run parameter. The accuracy of the system pressure measurement is high. But the deviation of the measured system pressure from the defined system pressure can reach ± 0.2 bar. It is mainly due to the non-stationary character of reflooding and a result of pressure regulators functioning. Considering the measured pressure curve instead of a constant value could improve test run simulation. But in the GRS sensitivity calculation it has been found that even variation of system pressure by 0.2 bar has a small effect on the results in comparison with the influence of the model uncertainties.

The inlet velocity of single-phase water (flooding velocity in experiment description) can be measured with high accuracy of about $\pm 1\%$ – $\pm 2\%$. In some test runs during the initial phase of the experiment, deviations from the declared test parameters are observed. If the deviation is large, for instance like in the test no. 233 (FEBA Series II), the measured parameters, instead of nominal values, should be considered in the modelling of the boundary condition.

Since the presented experimental uncertainties are not obtained from the experimenters, but rather as a result of the Phase III coordinator (GRS) estimation, an attempt to compare the FEBA experimental uncertainties with those of PERICLES experiment (Deruaz et al., 1985) has been performed. The estimated uncertainties of FEBA measurements are generally similar or larger than those of PERICLES. The uncertainties of measurements are mainly larger during the initial period of the test runs – when the flooding is initiated after the stationary heating phase of test section. As the experimental uncertainties were only estimations and not uncertainties defined by the experimenters,

the participants could correct them and consider them in the quantification process, if they deem it appropriate.

5.4. Results of model uncertainties quantification

The uncertainty ranges found by participants are compared for each type of physical models. Analysing the uncertainty ranges obtained by the participants it should be kept in mind that in many cases the quantified uncertainties are related to different codes and models. Theoretically, direct quantitative comparison should be limited to the users of the same code. One can also remind that, the uncertainties quantified with CIRCÉ should be regarded as a whole set of values, i.e. the value of calibrated parameter and its variation range are valid only together with the values obtained for the other parameter(s) considered in the quantification process and not alone.

In Tables 6–8 the uncertainty ranges of model specific uncertainties are compared. In Table 6 the uncertainties of wall heat transfer at dry-out conditions arranged according to the used method are listed. Majority of participants quantified HTC correlations according to the heat transfer regimes. Only two participants used global HTC as an uncertain input parameter. Two other used general HTC for dry out condition. KINS used the criterion for minimum film boiling temperature (MFBT) as an uncertain parameter, instead of a heat transfer coefficient. Some participants considered uncertainties of more than one wall heat transfer correlations. Few participants considered code specific parameters like wall heat transfer enhancement at the quench front and wall heat transfer enhancement due to the grid spacers. The coefficients related to heat transfer enhancement due to the grid spacers were used by KAERI to express the uncertainty of wall heat transfer in the range of steam convection. During the Phase II analyses, KAERI has found the multiplication factor for the heat transfer enhancement at grid spacers to be more influential than the one for steam convection. With the aim of not increasing too much the number of parameters to be quantified, KAERI selected only the parameter related to heat transfer enhancement for quantification. Taking it into account, the HTC uncertainty due to grid spacers as applied by KAERI has been considered as comparable with HTC parameters related to steam convection and included in the tables for comparison. By another hand, KIT who also used parameter related to the heat transfer enhancement at grid spacers, considered also the uncertainty of HTC from wall to vapour. In such configuration the grid spacer effect is separated from the HTC for steam convection and not really comparable with uncertainties of wall heat transfer related parameters. Indeed, the uncertainty range obtained by KIT for grid spacer heat transfer enhancement (0.0 – 1.1), is clearly out of the range of typical HTC uncertainties. For this reason, the parameter related to the heat transfer enhancement at grid spacers as used by KIT was not included in the comparison.

The parameters related to wall heat transfer enhancement at the quench front are very model specific and they are not compared in the tables. Even separate comparison of uncertainties of HTC for the quench front is not meaningful, since the parameters are related to different effects. In the case of GRS code ATHLET, the heat transfer enhancement at the quench front is considered only by the quench front progression but not by the wall temperature determination.

The criterion for minimum film boiling temperature used by KINS was included in all comparisons, since it is the only parameter for wall heat transfer considered by KINS. The minimum and the maximum of the ranges of the wall HTC are presented in tables in the following convention: {HTC – dry-out or film wall to gas correlations}/{HTC – global or film wall to liquid}. In parenthesis are in tables values related to wall dry/wet transition criterion.

The uncertain parameters related to the interfacial heat transfer model are listed in Table 7. Majority of participants selected as influential model uncertainty of interfacial heat transfer at the mist (dis-

persed droplet) flow. This is a typical flow pattern above the quench front. Some participants considered as uncertain parameter multiplication factor for global interfacial heat transfer. Parameter used by GRS is not an interfacial heat transfer multiplier but a number of droplets in the evaporation model. However, it can be recalculated on the basis of evaporation correlations used in the code ATHLET as a value equivalent with multipliers used by other participants. The recalculated value of this parameter is given in Table 7 in parenthesis. The model uncertainties related to momentum equation closure relations are listed in Table 8. The list of considered uncertain parameters is clearly dominated by multiplication factor for global interfacial friction. This parameter was considered by 8 participants. Five participants considered interfacial friction for mist flow. In addition to the interfacial friction some participants considered an additional parameter related to the phase relative velocity. These additional parameters are the droplet diameter in the case of two RELAP users, entrainment rate and wall friction. GRS used as parameter the multiplication factor for phase relative velocity. The uncertainty of phase relative velocity cannot be directly compared with uncertainty of interfacial friction factor. However, on the basis of relative velocity uncertainty an equivalent uncertainty of interfacial friction for the experiment condition could be roughly estimated.

Since almost all considered parameters are multiplication factors the best way to illustrate their ranges is to produce quotient Max (of the range) divided by the Min (of the range). This range presentation is given in Table 9. The terms related to the HTC are presented in Table 9 according to the same convention as in Table 6: {HTC – dry-out or film wall to gas correlations}/{HTC – global or film wall to liquid}.

The uncertainty ranges quantified by the participants differ considerably, even the uncertainties of the same codes. The best basis for comparison gives the uncertainties of heat transfer coefficients at dry out condition. This parameter has been considered by practically all participants. The comparison of the uncertainty ranges shows that they are quite similar for CATHARE code users and very different for RELAP users. More interesting appears the comparison according to the quantification method. The uncertainty ranges obtained by CIRCÉ users tend to small and moderate ranges. Uncertainty ranges obtained with FFTBM tend to large values. The methods of Tractebel and GRS seem to lead to moderate uncertainty ranges. Very large uncertainty ranges have been obtained by UNIPI and SJTU for film boiling and even larger for forced convection to gas by VTT. Such a large uncertainty range obtained by VTT for practically single-phase heat transfer correlation is astonishing. If obtaining of so large uncertainty range by VTT is related to application of FFTBM extended for consideration of more than one experimental run, is not clear. The possibility of consideration of more than one test is an obvious advantage. This enables to enlarge the data basis and take into account for instance tests considering different thermal-hydraulic conditions. It would enable to consider experiments performed in different test facilities. For quantification of model uncertainties, it is of importance to be able to consider many tests. Usually consideration of more experiments by uncertainties quantification leads to extension of the variation range. But it is a typical trend that a wider range of application is usually related with larger uncertainty of simulation. However, the cause of the very large uncertainty obtained by VTT remains unclear (extension of FFTBM, consideration of many tests). As it could be seen in the comparison, the FFTBM, applied even to only one experiment, results sometimes in quantification of large uncertainties. The reason for the large quantified uncertainty may be also the application of combination of CIRCÉ and FFTBM. In the opinion of the VTT the last reason is probably the most important one.

The direct comparison of the lower and upper limits of the quantified uncertainty ranges would be also of interest. For instance, for the HTC at dry out conditions the uncertainty ranges quantified for the RELAP code by CVRez and OKBM-1 are disjoint. The complete uncertainty

Table 9
Quantified uncertainty ranges (as max/min).

Participant	Code	Number of param	Wall HTC	Interfacial HTC	Momentum eqn – closure rel.	Method
CVRez	RELAP	2	/2.20	–	1.86	CIRCÉ
UPC	RELAP	3	/1.83	7.14	1.57	CIRCÉ
OKBM-1	RELAP	3	2.26/1.41	(1.37 – droplet)	–	CIRCÉ
Tractebel	RELAP	4	1.86/	2.75/2.75	4.86	IUQ
UniPisa	RELAP	5	8.0/2.6	18.0(3.80 – droplet)	1.86 (3.80 – droplet)	FFTBM
SJTU	RELAP	4	12.8/1.87	(2.61 – droplet)	3.0 (2.61 – droplet)	FFTBM
KIT	TRACE	6 + power	3.37/3.91	2.70	1.50	FFTBM
KINS	MARS-KS	2	(1.81)	13.3	–	CIRCÉ
KAERI-1	COBRA	4	/1.31	3.34	1.04	CIRCÉ
KAERI-2	COBRA	4	/1.57	2.18	3.04	MCDA
CEA	CATHARE	2 + 1(bias)	1.97/	–	1.05	CIRCÉ
IRSN	CATHARE	3	3.0/	–	73.	DIPE
Bel V	CATHARE	3	/2.56	–	1.04	CIRCÉ
GRS	ATHLET	6	1.47/2.0	3.2	~6.0 [2.5 – rel. velocity]	IUQ
VTI	APROS	6	34.0/(2.62)	7.0	4.29	FFTBM + CIRCÉ
OKBM-2	KORSAR	2	1.64/	–	4.92	CIRCÉ

range quantified by CVRez is below 1.0 and the range quantified by OKBM-1 above 1.0. But CVRez applied in their analyses global multiplier for all the heat transfer regimes, whereas OKBM-1 considered only a correlation for film boiling heat transfer regime. Moreover, OKBM-1 used a special code RELAP version SCDAPSI, what makes the comparison even more difficult. It appears that even for the heat transfer at dry out conditions, a phenomenon considered by all participants, only a qualitative comparison is possible.

The uncertainty ranges obtained for interfacial heat transfer are very different for RELAP code. No user of CATHARE code has considered it as uncertain parameter. Regarding the method of quantification, the uncertainty ranges obtained by participants using CIRCÉ as well as those using FFTBM are quite different. The ranges obtained by GRS and TRACTEBEL can be described as moderate.

The largest differences between uncertainty ranges can be observed for interfacial friction. The uncertainty ranges are very different for all codes. Some regularity can be found by uncertainty ranges arranged according to the quantification method. The participants using CIRCÉ quantified very small uncertainty ranges with exception of OKBM-2. The application of FFTBM lead also to rather small uncertainty ranges. However, two users of FFTBM used in addition to interfacial friction multiplier droplet diameter, a key parameter available in RELAP input and used in interfacial friction model. GRS and TRACTEBEL obtained moderate uncertainty ranges. IRSN using its own method DIPE obtained for the code CATHARE very large uncertainty range. It is particularly remarkable that the uncertainty range of interfacial friction obtained by IRSN for the code CATHARE was the largest uncertainty range among all uncertainty ranges obtained by participants. Simultaneously, the uncertainty ranges obtained for the code CATHARE interfacial friction by CEA and Bel V (users of CIRCÉ) were the smallest uncertainty ranges obtained by any participant in the Phase III. In the case of CEA, it is uncertainty of exactly the same model as that considered by IRSN.

The problem of the origins of these large differences appears to be very complex. CEA and Bel V did not consider pressure drops as responses in their quantification process, and this fact could lead to underestimation of the interfacial friction quantification. However, another CIRCÉ user OKBM-2, who did not considered pressure drops obtained quite reasonable ranges for interfacial friction. IRSN did not consider pressure drops as well. It seems that the main reason for the discrepancy would lie in the different assumptions of the both methods concerning the parameters dependency. The procedure resulting from

the assumption of parameter dependency leads to an extension of the uncertainty ranges.

The differences between CIRCÉ users are more difficult to explain. It may deal with the assumption of linear dependency between input and output parameters of the quantified model in CIRCÉ. It could be acceptable only for some codes.

5.5. Preliminary check of quantified model uncertainties

The last step in Phase III has been a preliminary check of the quantified uncertainty ranges, by performing an uncertainty and sensitivity analysis of the Test run 216 from Series I of FEBA. It is the FEBA test most similar to the PERICLES experiment, from the standpoint of boundary conditions. This preliminary check has been introduced for verification, if the applied procedure of quantification was successful.

The participants performed uncertainty analyses with the aim to obtain (95, 95) tolerance intervals for the following results:

- Cladding temperature as a function of time at three different levels: 3315 mm, 2225 mm and 1135 mm.
- Pressure drop along the total channel length and along the middle part of the channel
- Mass of water carried over the heated channel
- Quench front propagation

The results of the analysis were compared with the experimental data.

In general, the experimental data were bounded by the tolerance intervals. Only in the case of two CIRCÉ users and two FFTBM users some experimental data were not bounded by calculated (95, 95) uncertainty limits. It concerns mainly cladding temperatures at time when quench front is passing the level of thermocouple location. There were some discrepancies of experimental data and predicted uncertainty intervals at the beginning of the transient. The measured cladding temperature at 3315 mm decreases immediately after the start of reflooding, but the predictions showed instead a short period of temperature increase. The discrepancy can be attributed to differences between the reported initial conditions and the real ones during the test.

There are also discrepancies for the pressure drop along the total channel length. The measured value is clearly higher than the predicted ones at the start of the transient. The reason may be a difference between nominal and real initial conditions and/or a measurement error. Since the participants did not use the data measured at the very begin-

ning of the transient for quantification, it can be expected that it did not affect the quantified uncertainties.

For FEBA test 216, sensitivity analyses have been performed. Most of the participants identified as the most influential parameters for cladding temperature those related to the wall heat transfer at dry-out conditions. Some participants identified parameters related to interfacial heat transfer as the most influential for the cladding temperature.

A majority of participants identified the interfacial friction factor as the most influential parameter on pressure drop in the test section. GRS and Tractebel identified this factor as the most influential on all investigated output. Bel V identified interfacial friction factor as the most influential parameter not only for pressure drop but also for quench front propagation and cladding temperature in the upper part of the test section. Bel V determined very small variation range of the interfacial friction but a very large bias. Obviously, the interfacial friction factor was identified as so important parameter due to the large bias and not due to the variation range, which is one of the smallest one identified by participants.

As the most influential parameter for the water carried over, mainly interfacial friction factor and interfacial heat transfer coefficients have been identified.

CEA was the sole participant who identified parameters related to the special heat transfer model at the quench front as the most influential for quench front elevation, cladding temperature at the bottom of the test section and water carried over. This results from the large bias found by CEA for this parameter; however, the parameter was not varied explicitly. Other participants considering such parameters have not counted them among influential parameters.

Concerning quench front propagation, the majority of participants found the interfacial friction factor as the most influential one. Some found the HTC correlations as the most influential. Only CEA has found that the uncertain input parameter related to heat transfer enhancement at the quench front is the most influential parameter for the quench front propagation.

6. Phase IV: Verification and validation of the uncertainties found within Phase III

The goal of Phase IV has been the confirmation of the uncertainties determined during Phase III, by propagating them to the 6 FEBA tests considered in Phase III, and to the 6 tests of the 2D reflood PERICLES experiment. This latter step has been performed blindly.

Phase IV has been coordinated by CEA and IRSN. Fifteen organizations were involved (see Table 1). They are the same as involved in Phase III, with the addition of PSI. KIT had an incomplete participation in Phase IV, concerning only the uncertainty propagation for the FEBA tests.

The participants basically considered the same parameters as in Phase III, all of them related to physical models. Nevertheless, two participants suppressed one parameter each, because they were considered as included in another parameter. One participant (Tractebel) added 5 input parameters representing experimental uncertainties of the boundary conditions rather than model parameters (local heat flux, bundle power, inlet water temperature, system pressure and inlet velocity). Quantification performed by PSI represents conventional experts' judgement based approach.

PSI (not involved in Phase III) considered a large number of input parameters, including boundary conditions, material properties and physical models. The uncertainties of these input parameters were estimated by expert judgment, literature review and confirmatory uncertainty quantification based on the 6 FEBA tests available through PREMIUM.

In the specification of Phase IV it was required, both for FEBA and PERICLES, 200 code runs and the use of 5th and 195th order statistics, which are point estimators of the 2.5 and 97.5 percentiles, respectively.

They define a two-sided interval that has an expected coverage of 0.945. Some participants did not follow these specifications and used different number of calculations as well as different order statistics. The influence of these different choices is assumed to be of secondary order compared to the effect of the used quantification method and other, mainly the user, effects.

Phase IV has been developed in two steps:

- Step 1 (verification): uncertainty analysis of the 6 FEBA tests considered in Phase III.
- Step 2 (validation): uncertainty analysis of the 6 PERICLES tests. This step was performed blindly.

6.1. Uncertainty analysis of FEBA experiments

The analysis of the results has been performed according to two methods. The first one is qualitative and was carried out by CEA, the second one is quantitative and was carried out by IRSN. The outputs of interest for both kinds of studies are not the same ones.

The qualitative analysis performed by CEA is intuitive. To begin with, all the contributions are systematically analysed. More precisely, three issues are addressed for each figure:

1. Does the uncertainty band envelop the experimental time trend?
2. How the nominal calculation is located with respect to the experiment (under or overestimation)?
3. How wide is the uncertainty band?

For the CIRCÉ users, a fourth question must be raised:

4. Does the calibrated calculation improve the nominal calculation?

The results of this analysis were afterwards gathered in order to distinguish main trends and try to draw some conclusions.

This method was extensively applied to time trends of clad temperature:

- Clad temperature time trends at the middle of the bundle (respectively at 2225 mm and 1828 mm for FEBA and PERICLES);
- Clad temperature time trends in the upper part of the bundle (respectively at 1135 mm and 2998 mm for FEBA and PERICLES, knowing that for FEBA the elevation 0 mm corresponds to the top of the bundle, whereas for PERICLES, the elevation 0 mm corresponds to the bottom of the bundle);

Middle pressure drops time trends were also analysed, but less detailed, due to their very oscillatory behaviour.

For the quantitative analysis performed by IRSN, scalar outputs had to be provided by the participants. They were deduced from the time trends used by CEA for its qualitative analysis and are:

- For FEBA: The clad temperatures at the same time for all the time trends of a given test: the time of the maximum experimental clad temperature. These temperatures are considered in the middle and in the upper part of the bundle, i.e. respectively at 2225 mm and 1135 mm;
- For PERICLES: The maximum clad temperatures, even if the time where this maximum value is reached is not the same one for all the time trends. These temperatures are considered in the middle and in the upper part of the bundle, i.e. respectively at 1828 mm and 2998 mm;
- For FEBA and PERICLES: The quench times in the middle and in the upper part of the bundle, i.e. respectively at 1828 mm and 2998 mm.

6.1.1. Qualitative analysis

A synthesis of the CEA analysis is presented in Table 10. Participants have been ranked in four groups, according to the position of the uncertainty bands with respect to the experimental data of both types: cladding temperatures during the whole transient (not only the Peak Cladding Temperature, PCT) and quench times:

- First group (7 participants): the uncertainty bands envelop the experimental clad temperatures and quench times for all tests during the whole (or almost the whole) transient.
- Second group (5 participants): experimental data are bounded except for some tests just before the quenching.
- Third group (2 participants): experimental data are not enveloped during a rather long interval at the beginning of the transient
- Fourth group (3 participants): experimental data are systematically not bounded at the end of the transient.

The under or overestimation is indicated for the following scalar outputs:

- Clad temperature and quench time at 2225 mm
- Clad temperature and quench time at 1135 mm

The empty boxes correspond to cases where no systematic trend is observed or where the nominal calculations are globally good. In the last column, the width of the uncertainty bands is indicated, with 4 levels: very wide, wide, medium and narrow.

A main conclusion is that, except for the last group of participants, the envelop calculations are globally successful. It seems logical, since the participants used data from FEBA in the quantification, and the exercise was not blind. Only KAERI and KINS, among CIRCÉ users, considered exclusively the cladding temperature as responses, and excluded the quench times; as a consequence, their results do not envelop

the quench times. Moreover, KINS consider only 2 input parameters, without any parameter related to interfacial friction.

The 7 participants of the first group who enveloped all the data obtained very wide or wide uncertainty bands (except for CVRez). In the opposite side, the participants of the last group who did not envelop the experimental values had narrow bands. The two intermediate groups had mostly medium uncertainty bands. This means that the apparently more efficient or precise results failed to encompass the real results, and, conversely, those that did envelop were the less informative.

Another observation concerns the quality of the nominal calculations. Results may envelop real data even if the nominal calculations are not very satisfactory. It is very evident for the participants of the 1st group, whose bands are very wide.

In some cases, wide bands do not envelop data due to problems in the nominal calculation. The participants of the 3rd group have problems in predicting the adiabatic rise. On the opposite, there are participants who, despite having fairly narrow bands, envelop the data, because they obtain a good nominal calculation.

There is a very clear relationship between the type of quantification method and the width of the uncertainty bands. CIRCÉ and MCDA produced the narrowest bands. The widest bands were produced by PSI "method, IRSN method and SITU using FFTM.

A particularity of CIRCÉ is the possibility to estimate the so-called "calibrated value" of the input parameters, which is the median of their uncertainty distribution, in general different from the nominal value. Theoretically, the calibrated calculation improves the nominal one, by decreasing any systematic under or overestimation. All the CIRCÉ users performed this calculation, except for OKBM-KORSAR. For most of the participants, the calibrated calculation slightly improved the nominal calculation.

About the number of input parameters considered in the quantification process, it ranges from 2 (CVRez, OKBM-KORSAR, KINS) to 8 (Tractebel, who considered 3 inputs for Phase III, and 5 additional parameters for Phase IV). PSI considered the largest number, 26 input parameters, but used expert judgment in the quantification. The CIRCÉ

Table 10
Summary of the FEBA uncertainty results.

General result	Participant	Code	Method	Features of the nominal calculation				Bands width
				Tclad at 2225 mm	tquench 2225 mm	Tclad at 1135 mm	Tquench 1135 mm	
Exp. data bounded for all time trends	IRSN	CATHARE	DIPE			Underest.		Very wide
	SJTU	RELAP	IPREM	Underest.	Underest.	Underest.	Underest.	Very wide
	PSI	TRACE	Expert		Overest.			Very wide
	UNIPI	RELAP	IPREM	Underest.	Underest.	Underest.	Underest.	Wide
	VTT	APROS	FFTBM /CIRCE	Underest.		Underest.		Wide
	Tractebel	RELAP	IUQ	Underest.		Underest.	Underest.	Wide
	CVRez	RELAP	CIRCE	Underest.	Underest.	Underest.	Underest.	Medium
Exp. data bounded except before quench	CEA	CATHARE	CIRCE		Underest.		Underest.	Medium
	UPC	RELAP	CIRCE				Underest.	Medium
	OKBM	KORSAR	CIRCE		Unserest.		Underest.	Medium
	Bel V	CATHARE	CIRCE			Underest.		Medium
	GRS	ATHLET	IUQ	Overest.	Underest.	Overest.	Underest.	Wide
Exp. data not bounded at the beginning	OKBM	RELAP	CIRCE			Overest.		Narrow
	KIT	TRACE	IPREM			Overest.	Underest.	Medium
Exp. data not bounded	KAERI	COBRA	CIRCE		Underest.		Underest.	Narrow
	KAERI	COBRA	MCDA		Underest.		Underest.	Narrow
	KINS	MARS	CIRCE		Underest.		Underest.	Narrow

users consider generally less inputs than FFTBM and other methods. CIRCÉ estimates the range of variation of the input parameters so that the coverage of the experimental response by the uncertainty bounds is independent of the number of parameters. If the number of inputs is reduced, the ranges of variation increase in order to account for the total uncertainty.

The uncertainty bands for quench times and clad temperatures are wider at 1135 mm than at 2225 mm.

Another feature observed in the results is the influence of boundary conditions on the results. UNIFI noticed that the PCT prediction was better for tests with higher inlet flow rate, while the width of uncertainty bands seemed unchanged. On contrary, pressure had no influence on the quality of PCT prediction, but the width of the bands decreased when pressure increases. This behaviour illustrates the problems of extrapolating the uncertainty determined with a set of boundary conditions to different applications.

In the quantification of input uncertainties, some participants did not consider the quenching times as responses. The consequence was that, contrary to clad temperatures, quenching times were not enveloped. The opposite was also observed in the 3rd group of Table 10, the quench times were enveloped, but not the clad temperature. The reason seems to be a problem in the prediction of the adiabatic rise.

6.1.2. Quantitative analysis

The results of FEBA calculations were quantitatively analysed with IRSN methodology (Destercke and Chojnacki, 2008), using the IRSN SUNSET software (Chojnacki and Ounsy, 1996). Fifteen contributions of 24 scalar outputs (12 related to clad temperatures and 12 to quench times) were collected.

First of all, the percentage of uncertainty bands covering the corresponding experimental value was calculated, taking into account all participants. The average percentage is 83%, lower than the values of expected coverage for calculated tolerance bands. The frequency of coverage was better for clad temperatures (88%) than for quench times (77%).

The information provided to IRSN for each scalar output was an interval [lower uncertainty bound (LUB), upper uncertainty bound (UUB)] and a reference value (RV). Some participants did not provide the whole information for all outputs. For most of them, it was due to not attained UUB. Therefore, in order to perform the analysis on the same number of outputs, the not attained UUBs were set to the maximum of the UUBs given by participants. The IRSN methodology was

applied, and for each participant, two scores were calculated, quantifying the informativeness and the calibration (Fig. 7).

A negative correlation coefficient (-0.6) is found in these data, pointing out that wide uncertainty bands tend to bound the experimental data, giving a good calibration score, but they are not very informative. Conversely, narrow bands are very informative, but tend to fail in enveloping the data. UNIFI score is the closest to the centre of gravity of the cloud. KINS and KAERI, who are CIRCÉ users, provide narrow uncertainty bands and poor results that never encompass the experimental value for all time variables.

The effect of the quantification methodology has been studied as well. In Fig. 7 the results obtained by CIRCÉ are plotted in green. It is observed that the results are method-dependent. CIRCÉ produces narrow uncertainty bands, so that the informativeness score is generally high but the calibration score tends to be lower compared to other participants.

Figs. 8 and 9 show the calibration vs. informativeness scores plot separating the clad temperatures and the quench times. The very low calibration score of GRS, KAERI and KINS in Fig. 9 has an explanation: these participants did not consider as responses the quench times in their quantification process.

6.2. Uncertainty analysis of PERICLES experiments

The second part of Phase IV was the propagation of the model uncertainties obtained in Phase III to the results of the PERICLES tests. The exercise was performed blindly. As described in section 2.2, the difference of 2D reflood PERICLES experiment with respect to FEBA is the presence of 2D effects. The participants used different modelling criteria. In general, a multi-channel model with crossflows was chosen (e.g. RELAP users). CATHARE users applied 3-D model with exception of IRSN. IRSN used CATHARE with the 1D modelling, arguing that this would allow the extrapolation to PERICLES of the uncertainties found for FEBA. Using a 2D or 3D model would add a new set of uncertain models (mainly those governing the crossflows), which cannot be estimated from the FEBA experimental data.

The number of axial nodes in the heated part of the bundle is generally close to the corresponding number for FEBA, except for the users of a 3D model, who assigned a number of axial nodes for PERICLES significantly lower than for FEBA. On the contrary, PSI duplicated the number of nodes in moving from FEBA to PERICLES. The majority of participants model also the bottom plate, upper tie plate, upper core

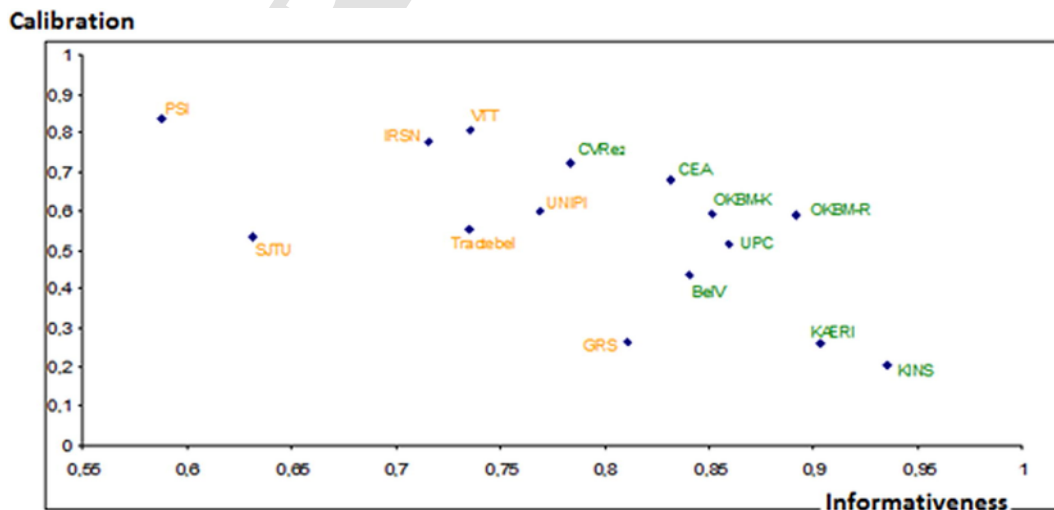


Fig. 7. FEBA: informativeness and calibration scores for each participant. Green color stands for CIRCÉ users.

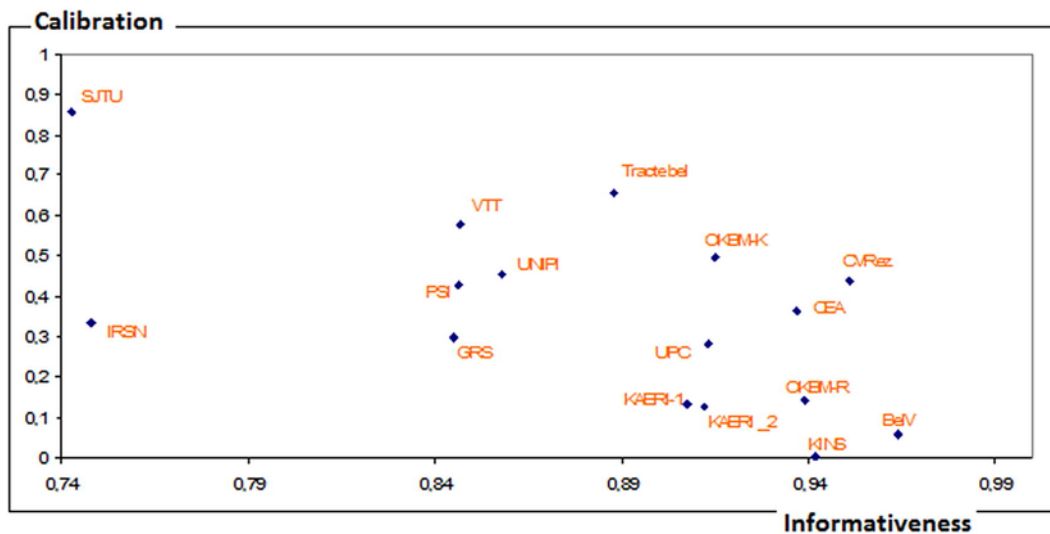


Fig. 8. FEBA: Informativeness and calibration scores for each participant for temperature variables.

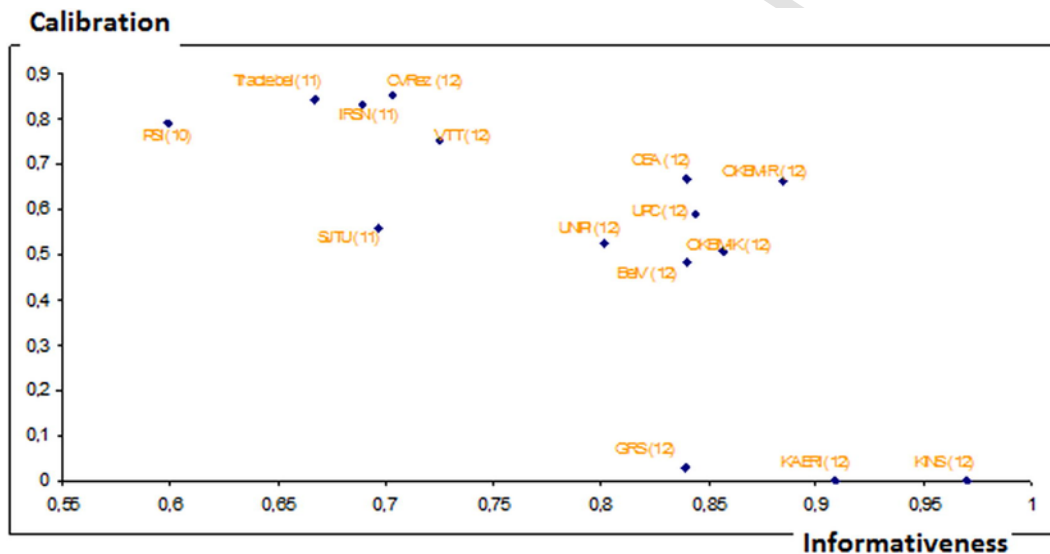


Fig. 9. FEBA: Informativeness and calibration scores for each participant for time variables.

plate and the housing. Typical average values were used for thermal conductivity and heat capacity of structural materials.

Unlike for FEBA, the experimental clad temperature profile was not given, except for the central assembly of one of the tests. The majority of participants followed by simulations as much as possible the experimental procedure, where the whole power was imposed to the fuel rods until the initial maximum clad temperature of the central assembly reached the specified value, and afterward starts the reflooding by the injection of cold water.

The theoretical maximum initial clad temperature is slightly higher than the really observed one. The reason is that the measurement of only one thermocouple was considered. This little discrepancy will be kept in memory for the comparison of the experimental clad temperatures with the uncertainty bands. It can explain why these uncertainty bands do not perfectly envelop the experimental data at the beginning of the transient, independently from the input uncertainties.

The quantified input parameters considered are the same as for FEBA with two exceptions:

- GRS considers 2 additional parameters: relative velocity in cross connection and bundle total power.

- PSI considers a total of 34 input parameters for PERICLES (26 for FEBA), with the 8 supplementary parameters being related to an increased number of boundary conditions for outlet pressure, inlet mass flow and temperature of the coolant, and heater rods power (physical model parameters were identical between the PERICLES and FEBA exercises).

The uncertainty propagation for PERICLES was based on 200 code runs, as required in the specification. Exceptions were IRSN (119 runs) and SJTU (93 runs). All the code runs were successful, except for PSI and Tractebel, who simply replaced the failed code runs by new runs.

6.2.1. Qualitative analysis

As for FEBA, CEA performed a qualitative analysis of the PERICLES results. In Table 11 a synthesis of results is provided, with the same structure than Table 10 for FEBA, but only considering three groups:

- First group (6 participants): the uncertainty bands satisfactorily envelop the experimental data for all time trends, even if they are not perfect at the beginning. The problem at the beginning is not induced by the input uncertainties but probably due to the lack of pre-

Table 11
Summary of the PERICLES uncertainty results.

General results	Participant	Code	Method	Features of nominal calculation				Band width
				Tclad at 1828 mm	tquench 1828 mm	Tclad at 2998 mm	tquench 2998 mm	
Exp. Data bounded for all time trends	VTT	APROS	CIRCE + FFTBM	Underest.	Overest.	Underest.	Very overest.	Very wide
	IRSN	CATHARE	DIPE	Overest.				Very wide
	SJTU	RELAP	IPREM	Underest.			Very overest.	Very wide
	UNIFI	RELAP	IPREM	Underest.	Underest.		Underest.	Wide
	PSI	TRACE	Expert	Underest				Very wide
	Tractebel	RELAP	IUQ	Underest		Underest.		Very wide
Exp. Data not always bounded	OKBM	KORSAR	CIRCE		Overest.		Very overest.	Medium
	UPC	RELAP	CIRCE		Underest.		Very underest.	Medium
	CVRez	RELAP	CIRCE				Very underest.	Medium
Exp. Data not bounded	GRS	ATHLET	IUQ		Very overest.	Underest.	Very overest.	Wide
	OKBM	RELAP	CIRCE			Very overest.	Overest.	Narrow
	KAERI	COBRA	MCDA		Underest.	Very overest.		Medium
	KAERI	COBRA	CIRCE		Underest.	Very overest.		Medium
	KINS	MARS	CIRCE		Underest.	Overest.	Very underest.	Narrow
	CEA	CATHARE	CIRCE	Very overest	Overest.		Overest.	Medium
	Bel V	CATHARE	CIRCE	Very overest.	Overest.		Very overest.	Narrow

cision in initial clad temperature. Uncertainty bands are wide or very wide.

- Second group (3 participants): experimental data are bounded most of the time, except for the end of the transient for some cases. Uncertainty bands are medium
- Third group (7 participants): experimental data are, in general, not bounded. Uncertainty bands are narrow or medium, with the exception or GRS (wide)

In the three groups, participants have been ranked by decreasing order of quality. As in the FEBA case, the table includes a classification of the systematic under or overestimations of the results, and of the uncertainty bands width.

More participants fail to envelop the experimental data for PERICLES than for FEBA. There is a relationship between the quality of nominal calculations and the bounding of experimental data. This relationship seems more decisive for PERICLES than for FEBA, especially when the quality is poor. All participants having unsatisfactory nominal calculations do not have successful envelop calculations, especially when the PCT is poorly predicted. The exceptions are participants with very wide uncertainty bands.

Nominal calculations are globally better in the first group, though in some cases the quench time is overestimated. In the second group, the opposite behaviour is observed, with underestimation of the quench time and thus a fail to envelop the data at the end of the transient. Participants of the 3rd group have poor nominal calculations, particularly for the prediction of PCT.

Participants gave miscellaneous reasons for the poor quality of the nominal calculations. For instance:

- Insufficiencies of the specifications, concerning the description of PERICLES. E.g. the housing and the heat losses through it.
- Deviation of the axial mesh centres from thermocouples position
- Deficiencies in the modelling of thermohydraulic phenomena: small break droplet breakup by spacer grids, wall-to-fluid heat transfer, interfacial friction, vapour mixing between central and lateral assemblies. The predicted reflood is, in some cases, too rapid, and too slow in others.
- Differences between the 1D and 3D modules of CATHARE.

As in the FEBA exercise, CIRCÉ and MCDA users have, in general, the narrowest bands. FFTBM and other methods users have wider bands, with similar widths. In the specific case of CIRCÉ users, no systematic change is observed when the FEBA-calibrated calculation is used.

Other remarks are the following:

- There are globally fewer input parameters considered in CIRCÉ. But, according to CEA, this fact does not theoretically have any impact on the width of uncertainty bands.
- It is generally more difficult to envelop the clad temperature in the upper part (2998 mm) than in the middle of the bundle (1828 mm). But in some cases, the opposite trend is observed. In FEBA, the bands

are wider at 2998 mm than at 1828 mm. This seems logical for quench times, but it is also generally found for clad temperature.

- Uncertainty bands are wider for the central assembly than for the lateral ones, at both elevations. It is normal, because clad temperatures and quench times are higher in the central assembly.
- There is a pressure effect: calculated results are degraded for the test at higher pressure, compared to the other PERICLES tests. Uncertainty bands are narrower when pressure increases.
- There are some cases where the 4 scalar outputs are bounded, but not the complete time curves. It is the case of participants having difficulties in the prediction of the adiabatic rise. Participants who did not consider quench time as responses in Phase III generally envelop the experimental quench times for PERICLES.

6.2.2. Quantitative analysis by IRSN

A total number of 36 responses, for different participants, were considered in the analysis (18 related to clad temperature, 18 related to quench time).

Firstly, the fraction of cases where the experimental value is enveloped by the uncertainty band was calculated. Percentages were between 37% and 87%, with a mean value of 65%, less satisfactory than in FEBA, and is far from the expected 95%. In average, the results are

better for quenching time than for cladding temperature (70% and 60%, respectively). In a number of cases the time variables did not attain upper uncertainty bounds, and the experimental value was considered to be enveloped as soon as it was larger than the lower endpoint of the provided interval (LUB).

Fig. 10 shows the plot of calibration vs. informativeness scores for the participants.

The values of the two criteria are more correlated (-0.8) than in FEBA. This means that narrow uncertainty bands are not able to envelop most of the PERICLES experimental values, and, therefore, that the extrapolation of uncertainties obtained for FEBA to PERICLES seems inadequate. The conflict estimator is 1 for all outputs indicating that participants are highly conflicting in their results.

The results are more method-dependent than the code dependent (see Fig. 11). The discrepancy between CIRCÉ and other methods is even more important than in FEBA. CIRCÉ produces narrow uncertainty bands encompassing few experimental values (high informativeness/low calibration). The other methods produce wide bands encompassing a large number of experimental data (low informativeness/high calibration). An analysis of variance reveals the inter-group variance to be ten times larger than the intra-group variance. The conflict indicator

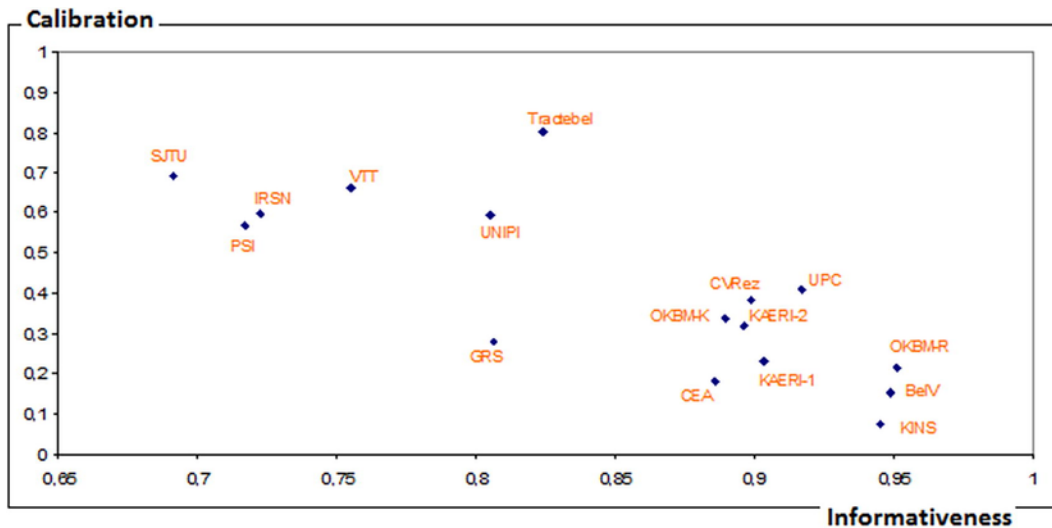


Fig. 10. PERICLES: Informativeness and calibration scores for each participant.

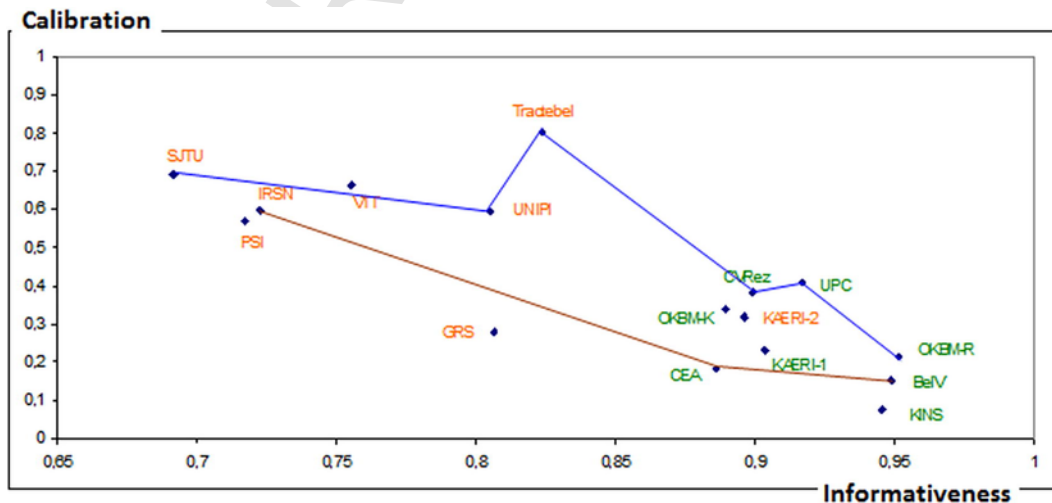


Fig. 11. PERICLES: Informativeness and calibration scores for each participant. Green color is for CIRCÉ users. The blue (resp. brown) solid line connects the participants using RELAP (resp. CATHARE).

is 1 for all outputs in the CIRCÉ subgroup and 0.73 for the other one. In both cases, it is higher than in FEBA.

Figs. 12 and 13 show informativeness and calibration scores, distinguishing temperature and time variables.

6.3. Axial pressure drops

The participants in Phase IV were also requested to provide 2.5% and 97.5 percentiles for the time trends of middle pressure drops. Only three participants included these responses in the quantification of model uncertainty. Interfacial friction is the most influential physical model for pressure drops; it was considered by majority of participants. CEA considered only the interfacial friction downstream from the quench front.

In the FEBA case, a lot of oscillations are observed in the time trends. For FEBA, the majority of participants envelop experimental data during the passage of the quench front, consistently with the results of cladding temperatures. At the end of the transient, the test section is filled up with liquid, so that the pressure drop is roughly the hydrostatic pressure. The participants did not reproduce well this value, probably because the pressure was calculated in different points than the measurements and because of the possible presence of vapour in

the calculation. The band widths are very variable among participants (from 0.5 to 4 bars).

Table 12 shows the ranking of participants according to their results. Two groups are clearly distinguished: those who bound the data for all time trends, and those who generally do not bound the data. The oscillatory character of the nominal calculation is described.

Participants who enveloped the clad temperatures were also successful with pressure drops. Those in the 2nd group had a bad nominal calculation. In some cases, the discrepancies between calculation and experiment are too high. In PERICLES, there was a two-phase mixture under the quench front at the end of the transient. A bad calculation of void fraction leads to a poor prediction of the pressure drop. The void fractions are controlled by interfacial friction upstream of the quench front.

The fact of taking into account the pressure drops as responses during Phase III does not necessarily imply successful envelop calculations. The quality of the nominal calculation and the width of the uncertainty bands seems to be more important.

7. Discussion of the results.

The main outcome of the PREMIUM benchmark was that the results of model uncertainties quantification are strongly related to the quan-

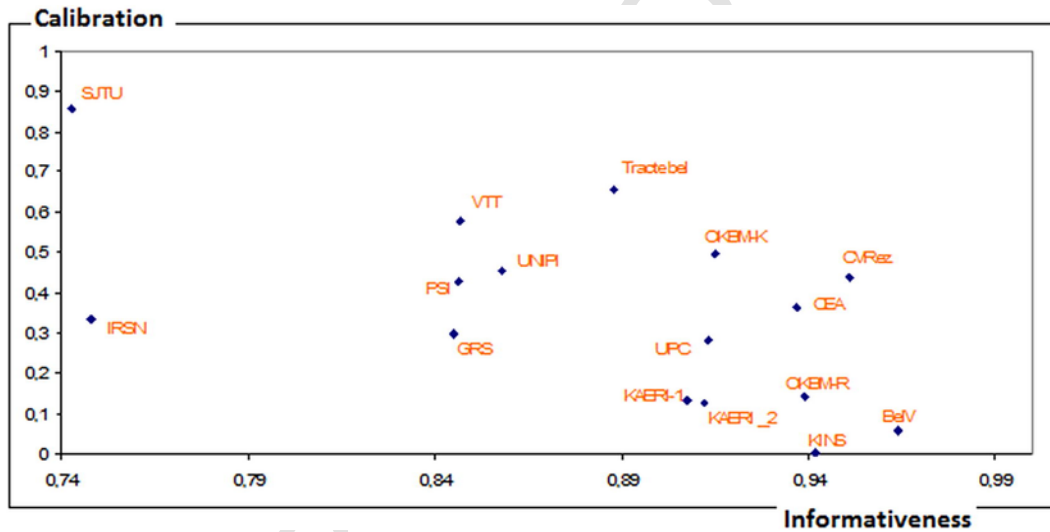


Fig. 12. PERICLES: Informativeness and calibration scores for temperature variables.

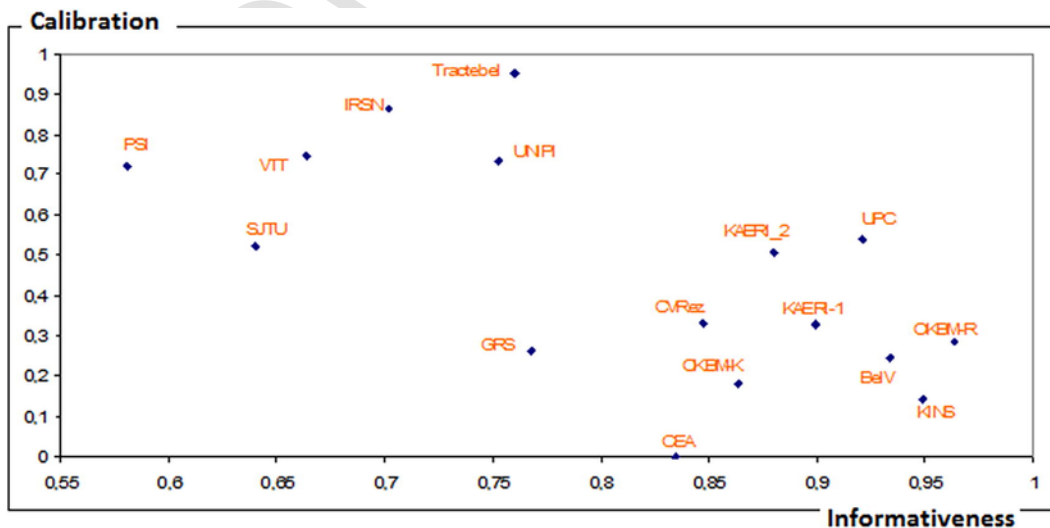


Fig. 13. PERICLES: Informativeness and calibration scores for time variables.

Table 12
PERICLES: Summary of the uncertainty results for the middle pressure drops.

General result	Participant	Code	Method	Nominal calculation – feature		Band width
				At quench front	At the end	
Exp. Data bounded	CVRez	RELAP	CIRCE	Very oscillatory	Very oscillatory	Medium
	IRSN	CATHARE	DIPE			Wide
	PSI	TRACE	Expert	Very oscillatory	Very oscillatory	Wide
	SJTU	RELAP	IPREM			Medium
	Tractebel	RELAP	IUQ	Very oscillatory	Very oscillatory	Medium
	UNIPi	RELAP	IPREM	Very oscillatory	Very oscillatory	Medium
	VTT	APROS	CIRCE/ IPREM			Medium
Exp. Data not bounded	Bel V	CATHARE	CIRCE	Underest.	Underest.	Medium
	CEA	CATHARE	CIRCE	Underest.	Underest.	Narrow
	GRS	ATHLEt	IUQ	Underest.	Underest.	Medium
	KAERI	COBRA	MCDA	Overest.	Overest.	Medium
	KAERI	COBRA	CIRCE	Overest.	Overest.	Medium
	KINS	MARS	CIRCE	Overest.	Overest.	Narrow
	OKBM	KORSAR	CIRCE	Underest.	Underest.	Medium
	OKBM	RELAP	CIRCE	Very oscillatory	Very oscillatory	Wide
	UPC	RELAP	CIRCE	Underest.	Underest.	Medium

tification method and are almost independent of the code used. The analysis of the Phase IV results showed that CIRCE produced narrower uncertainty bands than the other methods used in PREMIUM. VTT who applied CIRCE for calibration and FFTBM for uncertainty range quantification obtained wide and sometimes very wide uncertainty intervals. CIRCE users enveloped the experimental data for FEBA, while for PERICLES they generally failed to envelop the data. MCDA method gave similar results. The widest uncertainty intervals were obtained by PSI (expert judgement), SITU (IPREM) and IRSN (DIPE). Other IPREM users and users of other methods obtained wide or medium uncertainty intervals. The participants who obtained very wide uncertainty ranges enveloped experimental data.

A user effect is observed in the quantification of uncertainties. Participants using the same method and the same version of the same system code obtained significantly different uncertainty bounds. Depending on the kind of physical model uncertainty ranges obtained with different methods were different. For instance, the largest uncertainties of wall HTC were quantified by IPREM users and VTT, for interfacial HTC by one user of IPREM (Uni Pisa). Astonishing large uncertainty range was obtained by a CIRCE user (KINS). Extremely large uncertainty range for interfacial friction was obtained with DIPE method. On contrary, the CIRCE users obtained extremely small uncertainty ranges, partially not physical (interfacial friction below possible accuracy of experimental data) but CIRCE user, OKBM quantified large interfacial friction uncertainty. The IPREM user SITU, who produced along with PSI the largest uncertainty intervals, obtained continuously large, but not extremal, uncertainty ranges. As it can be seen from the performed uncertainty results, even the participants using the same code and the same quantification method obtained sometimes very different results. This may be due to differences in the input deck, which produce different nominal calculations, and to different choice of the input parameters and the responses.

The somehow unsatisfactory results of the benchmark PREMIUM resulted in the follow-up activities of some participants. The activities were related mainly to clarification of the reasons for poor results of

the quantification of model uncertainties and failure of fulfilling the validation step. Also, the observed tendency that quantified model uncertainties were dependent more on the quantification method than on the code, which uncertainties were quantified was quite confusing and a subject of further investigations.

7.1. Application of quantification methods

The results of the benchmark show that in the most cases the CIRCE users could not envelope the experimental data by uncertainty limits obtained during the validation step. The application of the calibration option in CIRCE was supposed to be the reason. It was of interest to clarify this problem. Therefore, the CEA has performed additional study, where quantification of model uncertainties with CIRCE were performed using the calibration option and not (Nouy and de Crecy, 2017). In fact, CEA has repeated the quantification of the physical models with the FEBA tests, with the same parameters and responses, but without estimation of a calibrated value. Once estimated the uncertainty for the input parameters, the propagation has been performed for both FEBA and PERICLES tests. The validation of the model uncertainties has shown much better results. In the case without calibration the experimental data were generally enveloped by the uncertainty limits. The results have clearly shown that application of the calibration was disadvantageous. For the reasons outlined, CEA changed the CIRCE users' guidelines and recommends the suppression of the recalibrated value estimation, especially when the database used for recalibration is small and does not cover the physical conditions in which the model will be used.

In the PREMIUM benchmark, KINS participated with two different methods of model uncertainties quantification. They used CIRCE method and their own MCDA method based on data assimilation. In both cases the validation step was not successful. The obtained uncertainty limits did not envelope the experimental data. The applied methods have similarities. Both methods use calibration of the nominal calculation in the quantification procedure. However, even though the

MCDE method considers also non-linear dependencies of uncertain parameters and calculation responses, no significant improvement of the quantification results is observed. It indicates that in this case the application of calibration is the reason of the disappointing results of the MCDE methodology application.

In the PREMIUM benchmark, UPC used the CIRCE method to quantify model uncertainties. After the Phase IV of the benchmark, UPC performed an additional analysis (Freixa et al., 2016) using the IPREM method. They consider the same model uncertainties as in the case of CIRCE but applied only to one test, the Test no 216 of FEBA Series I. For CIRCE application they used all 6 tests of the FEBA Series I. However, with the IPREM method, the quantified uncertainty ranges were wider than those obtained with CIRCE. The validation step performed with uncertainties obtained with the IPREM method was successful: the experimental data of PERICLES experiment were enveloped by uncertainty limits obtained by uncertainty analysis with model uncertainties quantified with IPREM. Whereas the validation step performed by UPC within the PREMIUM was not successful: the experimental data were partially not bounded by uncertainty limits. The results of this comparison indicate also that the calibration used in the course of uncertainties quantification is disadvantageous.

Since the validation step was not successful also either for GRS, the GRS performed a follow-up analysis of the benchmark calculations as well. However, even the uncertainty intervals obtained by GRS for the PERICLES test were quite wide, the experimental data were not bounded. In the follow-up analysis (Skorek, 2017) reasons for the failure of the validation step were found:

- Axial discretisation of the test section,
- Using of standard boron nitride material properties from ATHLET, instead of that defined in the experiment description
- Over-prediction of the entrainment rate
- Relative velocity modelling in the range of dispersed droplet flow.

Correction of the axial discretisation, use of the of boron nitride material properties as in the experiment description, and in addition application of different physical model in the range of dispersed droplet flow, produced a best estimate calculation agreeing very well with temperature measurements. Starting from this calculation as reference, the measured data would be bounded using quantified model uncertainties. In the verification step, GRS observed some discrepancy between experimental data and calculated uncertainty limits. Analysing FEBA simulations it has been found out that the axial power distribution was not axial-symmetric as it was claimed in the experiment description, but top skewed. Apparently, according to the dependency of heating material electrical resistance with the temperature, which changes during the experiment. Also, an unexpected sampling effect has been discovered. In the sample of randomly varied uncertain parameters, used for model uncertainties quantification, an exceptional combination of parameters occurred, which lead to extremal values of uncertainty limits. Such or similar combination was not generated in the sample used for verification procedure. This was the reason that in the verification step the experimental data were not completely enveloped by uncertainty limits, as it could be expected.

7.2. User effects

PREMIUM has shown that the quantification methods of model parameter uncertainty have a significant user effect. Their application needs the use of engineering judgment, in addition to the well-known user effect issue when using system codes.

The existence of user effect can be detected in the different stages of the application of a quantification method. The following features of the quantification, which depend on the judgment and experience of

the user, are important for quantification results:

- 1) Choice of the responses (outputs) on which the quantification is based
- 2) Choice of the model parameters to be quantified
- 3) Selection of the database parameters (experimental measurements) used in the quantification
- 4) Choice of input parameters in the modelling of experiments. This includes the selection of specific process models, coefficients and nodalisation approaches
- 5) Selection of options or setting criteria by application of quantification methods

The selection of model parameters to be quantified is very important. For the selection of these parameters, using sensitivity studies, the preliminary determination of the investigated parameters uncertainty is of importance. Regarding for instance, the interfacial friction coefficient, (see Fig. 5 and Table 8), it can be seen that in many cases preliminary estimation differs strongly from the quantified values. Few participants quantified similar, to the preliminary estimated, uncertainty ranges. Only IRSN quantified larger uncertainty range than it was estimated. The majority of participants quantified smaller, sometimes much smaller, uncertainty ranges as initially assumed. Nevertheless, it can be estimated as favourable since, it is better to select a parameter which appears thereafter as not important, than neglect an influential one.

The model quantification can be viewed as a backpropagation of uncertainties, from the responses (outputs) to a specific set of inputs. In a sense, it may be said that the backpropagated uncertainty is apportioned among the selected inputs. Usually, input uncertain parameter express (in)accuracy of a particular physical model and should be mutually independent. Even though, the uncertainty quantified for a given parameter depends, in general, on the set of selected parameters.

The reflection of the details of a model is also of importance. Is it enough to express the uncertainty of the model with one parameter (e.g. multiplication factor) or use more specific parameters? But using more parameters requires additional detailed experimental data, which are not always available, to quantify all of them.

The choice of responses to be used in the quantification is also important. The very low calibration score of some participants shown in Fig. 7 is due to the fact that they did not consider as responses the quench times or pressure drops in their quantification process. Responses must be clearly dependent on the selected input parameters. But application of inverse methods for combined effect tests may require, for quantification of some phenomena, the use of responses that are not mutually dependent. In praxis the responses can be dependent on several quantified parameters. Special statistic methods are required to solve this problem. Responses should be used so that the effects of various parameters could be allocated to the right one, e.g. not selecting the pressure drop is an error, as the uncertainty of temperature simulation cannot be distributed to interfacial friction and heat transfer coefficient correctly.

7.3. Extrapolation of experimental database

In the course of the benchmark the question arose, if the FEBA experiment is a suitable basis for model uncertainties quantification for reflooding test facility like PERICLES.

There is a considerable difference in the treatment of the housing walls of both FEBA and PERICLES test bundles. The housing of the test facility should reproduce the influence of surrounding fuel rods (large array of fuel rods) as in reactor geometry. It is necessary to consider the test bundles as representative for reactor scale geometry. In the FEBA test facility there were thick walls heated initially from the bun-

dle and isolated to minimize the heat losses to environment. In PERICLES there were thin walls with heating to keep the wall temperature few degrees above the saturation level. It is difficult to estimate the influence of both types of housing on the simulated reflooding process.

A clear difference which may have even larger influence on the experimental results is the size of test bundles. In the FEBA test facility 16 rods, out of 25, are in the vicinity of housing walls. In the larger PERICLES test facility 112 rods out of 357 are near housing walls. It means majority of rods in FEBA test facility are peripheral rods and the reflooding of the test bundle may be considerably affected by housing walls, much stronger than in the PERICLES facility. It means the FEBA test facility is not really representative for large bundles like in the PERICLES experiment. As this fact can influence significantly the quantified model uncertainties, the extrapolation of the results can be disturbed by a distorted scaling between both facilities. This was not considered defining the benchmark, since in both experiments it was claimed that the test facilities are representative for reactor geometry.

8. Conclusions

As described in the previous section, results of model uncertainties quantification appeared to be more dependent on the quantification method than on the thermal-hydraulic code used. Large differences have been observed in the quantified uncertainty, depending on the used method but also depending on the user. This is an astonishing result, because what was expected is a strong dependency on the codes. The physical models, even if they describe the same phenomenon, are different and have a specific implementation in each system code. The benchmark PREMIUM enabled to review the actual state of praxis of model uncertainties quantification and to draw important conclusions concerning the necessary improvement of the applied procedures and approaches. However, the applied methods appeared as not matured for practical applications. The benchmark results were dominated by user effect. The application of all methods depended a lot on user expertise and inconsistency of application by different users could be observed.

As the main problems of the methods application could be identified:

- Using of calibration
- Qualification of input data for reference calculation
- Experimental basis for quantification
- Selection of adequate model parameters, measured values, and code responses for quantification procedure.

The follow-up activities of several participants enabled to solve some problems and clarify important aspects of application of the methods. In these studies, the identified problems were at least partially solved, and a satisfactory evaluation of model uncertainties could be achieved. It indicates that, when applied properly, the methodologies used in the benchmark are suitable tools for model uncertainties quantification.

Concerning the problems and improvements of model uncertainties quantification, important conclusions regarding the particular aspects of the model uncertainties quantification could be drawn.

8.1. Quantification database

A very important step is the definition of the database for development, validation and quantification of a physical model. The defined database determines the range of validity of the quantified uncertainties. The database used for model uncertainties quantification consists, first of all, of experimental data. However, in some cases, reference model or calculations, e.g. CFD calculations, can be applied as data for evaluation of simpler models.

A compromise must be found between general applicability and special application of the developed database. The database is set usually for special case application, in the sense that it is related to the foreseen application of the quantified uncertainties. In other words, if model uncertainties are needed for calculating a specific scenario in a plant, the database should include experiments related to this scenario. But the database can be also more general, so that the quantified model uncertainties are applicable to a wide spectrum of simulations. This is rather the case by quantification of model uncertainties in the frame of so called advanced (Unal et al., 2011) or extended (Skorek, 2018) code validation. By the uncertainty analyses performed by code users, the database and the following uncertainties quantification are usually for special application case.

The experiments included in the experimental database have to be adequate for the foreseen application. The experimental uncertainties should be carefully examined, because they can influence the quantification significantly.

8.2. Input data qualification

The quality of the nominal calculation is important for the results of the propagation calculations. A good nominal calculation is the basis of successful quantification of model uncertainties and its validation. Qualification of input data involves suitable nodalisation schema as well correct representation of facility features, initial and boundary conditions.

8.3. Calibration

The comparison of the CIRCE application using the calibration of the basic calculation with the application of nominal calculation confirmed the opinion that apply calibration in the quantification of model uncertainties is not advantageous. The problem of model calibration in the uncertainty analysis was controversially discussed during the Phase III. On the one hand, the calibration was recognised as a possibility of more precise quantification of model uncertainties. On the other side, it was regarded as a not acceptable modification of the code, which makes questionable the best estimate character of the code and its validation. In particular, it is inconsistent with the definition of the best estimate features when results of the best estimate code models are outside of the quantified uncertainty ranges. It can lead to the situation when best estimate reference calculation is outside of the uncertainty interval obtained in the course of uncertainty analysis. This is a clear inconsistency with the best estimate code definition.

8.4. Quantification procedure

In many instances, model parameters are adjustable coefficients used for fitting the models to experimental data and may have only limited physical meaning. These fitted models may have poor ability of extrapolation outside the range of development and validation field.

The quantified uncertainty obtained for a specific parameter strongly depends of the total set of simultaneously quantified parameters. This means that quantified uncertainties are attributes of the total set of parameters, as well as intrinsic properties of individual parameters. The correct allocation of total simulation uncertainty to individual parameters depends on the quantification method. In fact, it is the main task of the quantification method and simultaneously the most important criterion of the method applicability.

The set of quantified parameters must include the most influential ones on the responses; otherwise the resulting uncertainty may be completely misleading. On the other hand, it is advisable to include in the quantification all potentially important model parameters, not only the

most influential ones, because in other applications the set of dominant parameters may be different.

Quantification methods investigated in the benchmark PREMIUM are dedicated to combined effect tests. Nevertheless, some parameters related to the phenomenon may be quantified on the basis of SETs. In such case, it is important to have guidelines about how to proceed: using for quantification only the SET data or combine them with CETs data.

In a computational code, models are organized in a hierarchical structure, so that an individual model generally encompasses several sub-models or correlations. This structure has to be considered in the quantification of model parameters, and guidelines are needed for that matter. For instance, the different results obtained by quantifying different sub-models or by quantifying the complete model via a global multiplier should be analysed.

Methods used for model uncertainties quantification should not be applied at the same time to initial conditions, boundary conditions, material properties, and other magnitudes having full physical meaning, unless there is no other source of information about their uncertainty.

8.5. User effect

Complex physical models may have a considerable number of physical parameters and produce a large number of responses. The results of the quantification of model parameters uncertainty are very dependent on the selected parameters to be quantified and the selected responses to be used in the process. For the selection of the potentially important parameters, the use of sensitivity studies, the preliminary determination of the investigated parameters uncertainty is of importance. The estimated uncertainty ranges influence a lot the results of sensitivity studies and selection of parameters.

The selection of parameters, responses and database are fundamental parts of quantification methods. Guidelines and procedures should be established for such processes. Without these types of guidelines and procedures, the methods will have a strong user effect. Quantification methods are tools to minimise the engineering judgment, but they cannot completely eliminate it.

8.6. Final conclusion

The final conclusion of the benchmark is that a systematic approach is needed for comprehensive quantification of model uncertainties. Such a methodology has to consider the following steps, which have been identified as important for quantification procedure:

Specification of the problem, in particular identification of the models which uncertainty should and can be evaluated in the frame of the quantification procedure

Development of the experimental database

Selection and assessment of simulation model: selection of optimal available models for the intended task and qualification of the input data used for simulation of the experiments in the frame of the applied method

Application of the method for modelling uncertainties quantification

Validation of the results of the uncertainties quantification

Assessment of the predictive capability of the evaluated model uncertainties in the context of the intended application

The quantification of model uncertainties is a complex procedure and needs to be performed in a systematic way. It can be done by the code users in the course of a particular application. Even better would be, if the model uncertainty quantification is performed by the code developers in the frame of code validation and incorporated in the code documentation. The probably more general model uncertainties quantification done by code developers could be improved for intended application by code users, if they recognise it as necessary.

The direct results of the finding regarding necessity of a systematic approach was the launch of the follow-on project SAPIUM (Baccou et al., 2018). This project deals with the elaboration and the formulation of the good practice for a good practice for model uncertainties quantification.

Uncited reference

.

Acknowledgments

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

References

- Baccou, J., et al., 2018. SAPIUM: A systematic approach for Input uncertainty Quantification. Proceedings BEPU2018 Conference, Paper no 326.
- Cacuci, D.A., Ionescu-Bujor, M., 2010. Best-estimate model calibration and prediction through experimental data assimilation – I: mathematical framework. Nucl. Sci. Eng. 165, 18–45.
- Chojnacki, E., Ounsy, A., 1996. Description of the IPSN method for the uncertainty and sensitivity analysis and the associated software: SUNSET. Proc. ASME/JSME, ICONE, Louisiana, USA 3, 545–550.
- Deruaz, R., Clement, P., Vateau, J.M., 1985. Study on two-dimensional effects in the core of a light water reactor during the ECC's phase following a loss of coolant accident. Nucl. Sci. Technol. Series Rep. (EUR 10076 EN).
- de Crécy, A., 2001. Determination of the uncertainties of the constitutive relationships of the CATHARE 2 code. Proc. of Int. Conf. M&C 2001.
- Destercke, S., Chojnacki, E., 2008. Methods for the evaluation and synthesis of multiple sources of information applied to nuclear computer codes. Nucl. Eng. Des. 238 (9), 2484–2493.
- Freixa, J., de Alfonso, E., Reventos, F., 2016. Testing methodologies for quantifying physical models uncertainties. A comparative exercise using CIRCE and IPREM (FFTBM). NED 305, 653–665.
- Heo, L., Lee, S.-W., Kim, K.D., 2014. Implementation of Data Assimilation Methodology for physical Model Uncertainty Evaluation using post-CHF Experimental Data. Nucl. Eng. Technol. 46, 5.
- Ihle and Rust, 1984a. Ihle P., Rust K., 1984a. FEBA – Flooding Experiments with Blocked Arrays, Evaluation Report, KfK Karlsruhe, Rep. KfK 3657.
- Ihle and Rust, 1984b. Ihle P., Rust K., 1984b. FEBA – Flooding Experiments with Blocked Arrays, Data Report 1, Test Series I through IV, KfK Karlsruhe, Rep. KfK 3658.
- Joula, J., Probst, P., 2008. DIPE: determination of input parameters uncertainties. Methodology applied to CATHARE V2.5.1. J. Power Energy Syst. 2 (1), 409–420.
- Kovtonyuk et al., 2015. Kovtonyuk A., Petrucci A., D'Auria F., 2015. "Post-BE-MUSE Reflood Model Input Uncertainty Methods (PREMIUM) Benchmark. Phase II: Identification of influential parameters", NEA/CSNI/R(2014)14.
- Kovtonyuk, A., Lutsanych, S., Moretti, F., D'Auria, F., 2017. Development and assessment of a method for evaluating uncertainty of input parameters. NED 321, 219–229.
- Nuoy, E., de Crécy, A., 2017. Quantification of uncertainty of physical models integrated into system thermohydraulic codes. NED 321, 278–287.
- OECD/NEA/CSNI, 2011. OECD/NEA/CSNI, 2011. "BEMUSE Phase VI Report: status report on the area, classification of the methods, conclusions and recommendations". NEA/CSNI/R(2011)4.
- Skorek, T., 2017. Input uncertainties in uncertainty analyses of system codes: quantification of physical model uncertainties on the basis of CET. NED 321, 301–317.
- Skorek, T., 2018. V&V and Uncertainty Quantification of Code Models. Proceedings of ANS BEPU2018 International Conference, Paper no 110, Lucca, Italy.
- Unal, C., Williams, B., Hemez, F., Atamturktur, S.H., McClure, P., 2011. Improved best estimate plus uncertainty methodology, including advanced validation concepts, to license evolving nuclear reactors. Nucl. Eng. Des. 241, 1813–1833.
- Wilks, S.S., 1941. Determination of sample size for setting tolerance limits. Ann. Math. Statist. 12, 91–96.
- Zhang, J., Dethioux, A., Kovtonyuk, A., Schneidesch, C., 2019. Development of a pragmatic approach to model input uncertainty quantification for BEPU applications. Nucl. Technology 205, 140–152.