HORIZON 2020

HORIZON 2020 RESEARCH AND INNOVATION FRAMEWORK PROGRAMME OF THE EUROPEAN ATOMIC ENERGY COMMUNITY

Nuclear Fission and Radiation Protection 2018 (NFRP-2018-4)

Project acronym: **SANDA**

Project full title: **Solving Challenges in Nuclear Data for the Safety of European Nuclear facilities**

Grant Agreement no.: **H2020 Grant Agreement number: 847552**

Work package N°: ***WP5***

Identification N°: ***D5.6***

Type of document: ***Deliverable***

Title: REPORT ON CORRELATION BETWEEN INTEGRAL EXPERIMENTS

Dissemination Level: ***PU***

Reference:

Status: ***VERSION 1***

Comments:

	Name	Partner	Date	Signature
Prepared by:	V. Bécares	1		Firmado por BECARES PALACIOS VICENTE - DNI ***8188** el día 03/07/2024 con un certificado emitido por AC Sector Público
WP leader:	R. Jacqmin	3		JACQMIN Robert 116318 Signature numérique de JACQMIN Robert 116318 Date : 2024.07.03 23:21:39 +02'00'
IP Co-ordinator:	E. González	1		Signed by GONZALEZ ROMERO ENRIQUE MIGUEL - DNI ***3708** at 04/07/2024 with a certificate isued by AC Sector Público

SANDA Project D5.6: Report on correlation between integral experiments

V. Bécares[1], N. García-Herranz[2], O. Cabellos[2], A. Rizzo[3], D. Bernard[3], I. Kodeli[4], G. Žerovnik[4], A. Berger[4]

[1]*CIEMAT, Avenida Complutense, 40, 28040 Madrid (Spain).*

[2]*Universidad Politécnica de Madrid (UPM), José Gutiérrez Abascal, 2, 28006 Madrid (Spain).*

[3]*CEA, DES, IRESNE F-13108 Saint-Paul-Lez-Durance (France)*

[4]*IJS, Jamova 39, Ljubljana (Slovenia)*

Contact: vicente.becares@ciemat.es, Tel: +34 91 346 63 85.

Contents

Table of acronyms

ANL	Argonne National Laboratory (USA)
ANN	Artificial Neural Network
ASTRID	Advanced Sodium Technological Reactor for Industrial Demonstration
CEA	*Commissariat à l'Énergie Atomique et aux Énergies Alternatives* (Alternative Energies and Atomic Energy Commission, France)
CML	Critical Mass Laboratory (Rocky Flats, USA)
DICE	Database for ICSBEP
DOE	Department of Energy (USA)
EG UACSA	Expert Group on Uncertainty Analysis for Criticality Safety Assessment
ESFR	European Sodium Fast Reactor
HCF (HEU-COMP-FAST)	ICSBEP designation for reactors with Higly-Enriched-Uranium fuel in Compound form and with a Fast neutron spectrum
HST (HEU-SOL-THERM)	ICSBEP designation for reactors with Higly-Enriched-Uranium fuel in Solution form and with a Thermal neutron spectrum
ICSBEP	International Criticality Safety Benchmark Evaluation Project
IPEN	*Instituto de Pesquisas Energéticas e Nucleares* (Nuclear Energy Research Institute, Brazil)
IRPhE	International Reactor Physics Experiment Evaluation
IPPE	Institute of Physics and Power Engineering (Russia)
JAEA	Japan Atomic Energy Agency
KNN	K-Nearest Neighbours
LANL	Los Alamos National Laboratory (USA)
LLNL	Lawrence Livermore National Laboratory (USA)
LCT (LEU-COMP-THERM)	ICSBEP designation for reactors with Low-Enriched-Uranium fuel in Compund form and with a Thermal neutron spectrum
MCF (MIX-COMP-FAST)	ICSBEP designation for reactors with Mixed uranium and plutonium fuel in Compund form and with a Fast neutron spectrum
MCT (MIX-COMP-THERM)	ICSBEP designation for reactors with Mixed uranium and plutonium fuel in Compund form and with a Thermal neutron spectrum
MMF (MIX-MET-FAST)	ICSBEP designation for reactors with Mixed uranium and plutonium fuel in Metallic form and with a Thermal neutron spectrum
MCNP	Monte Carlo N-Particle
ML	Machine Learning
NEA	Nuclear Energy Agency of the OECD
OECD	Organization for Economic Cooperation and Development
ORNL	Oak Ridge National Laboratory (USA)
PMF (PU-MET-FAST)	ICSBEP designation for reactors with Plutonium fuel in Metallic form and with a Fast neutron spectrum

PNL	Pacific Northwest Laboratory (USA)
PST (PU-SOL-THERM)	ICSBEP designation for reactors with Plutonium fuel in Solution form and with a Thermal neutron spectrum
SANDA	Supplying Accurate Nuclear Data for energy and non-energy Applications
SCALE	Standardized Computer Analyses for Licensing Evaluation
SINBAD	Shielding Integral Benchmark Archive and Database
SFR	Sodium Fast Reactor
VNIIEF	All-Russian Scientific Research Institute of Experimental Physics
VNIITF	All-Russian Scientific Research Institute of Technical Physics
ZPPR	Zero Power Physics Reactor
ZPR	Zero Power Reactor

1 Introduction

The main purpose of SANDA project Task 5.2 is to assess the JEFF evaluated nuclear files using existing experiments, with the aim to contribute to the development of the future JEFF-4.0 nuclear data library. Calculations (C) are compared with experimental measurements (E) and results are analysed to identify possible biases and trends. However, while the comparison between C and E is essential and provides much information, it does not make the most of all the information provided by the experimental measurements. Adjustment and assimilation studies are powerful techniques to make the most of the information provided by experimental measurements and reducing biases in the calculation results and providing recommendations for nuclear data improvement.

Adjustment and assimilation techniques usually need to employ a large set of integral experiments to provide consistent information. However, the experimental parameters of the experiments within this set can be correlated, and for the application of these techniques it is essential to have some prior knowledge of these correlations. These correlations may arise from the use of the same facility, materials or measurement techniques, among other elements. To stress the importance of an adequate knowledge of these correlations, it is worth mentioning that the Working Party on International Nuclear Data Evaluation Co-operation (WPEC) of the OECD/NEA has very recently approved the proposal (May 2024) and established a new subgroup (as successor of the former subgroup 46) whose objectives include *"the determination of experimental correlations between integral benchmarks; and to assess their importance in a nuclear data adjustment"*.

In this line, subtask 5.2.1, entitled *"Assessing correlations in integral experiments"* has been included within SANDA project Task 5.2. More specifically, the description of task 5.2.1 reads:

Subtask 5.2.1: Assessing correlations in integral experiments
While a considerable effort has been given to nuclear data covariances in recent years, much less attention has been paid to correlations in integral experiments used in validation, adjustment, and assimilation studies. In point of fact, correlation coefficient data for criticality cases are available for only 93 integral experiments of the DICE database associated with the ICSBEP Handbook.
Although this project will not attempt to produce adjusted nuclear data libraries nor to assimilate validation information, ***CIEMAT, JSI, CEA/DEN, and UPM will share their best experts' opinions on the "missing correlations in integral experiments" problem, with the goal of assessing its impact on nuclear data validation studies. Simulations will be made to estimate the correlations between the experimental uncertainties of integral experiments and quantify their impact on some reactor concept.***

This report presents the results of this subtask. It is structured as follows. In section 2 and 3, the origin of the experimental correlations and their impact are described. Then, in section 4, a survey of the available correlations between integral experiments is presented. Although this study mainly focuses on criticality benchmark experiments, a review of the status of correlations in between shielding benchmarks is also included. In section 5, the methodologies used for calculating correlations between integral experiments are discussed. Finally, in section 6, four applications to particular cases are presented: (1) two cores of the EOLE facility investigated under the CAMELEON program, (2) a set of six cores loaded in the ZPR facility, which are considered to be of interest for validation of nuclear data for Sodium Fast Reactors (SFRs); (3) reaction rates measured at the ASPIS Iron-88 benchmark experiment, and (4) an example of the application of Machine Learning techniques in the interpretation of uncertainties and correlations in a benchmark experiment.

2 Origin of experimental correlations

Measured integral parameters of nuclear reactors such as the multiplication factor are always accompanied by experimental uncertainties. These uncertainties arise from uncertainties in the material compositions, impurities, densities, dimensions, etc. of the experimental facility, as well as in the measurements techniques used. These uncertainties are provided by the experimentalists or inferred from the experimental documentation and based on that, the benchmark evaluators propagate the uncertainties in physical parameters to obtain an experimental uncertainty in the integral responses. Although this procedure is a complex task, since experimental uncertainties are not always easy to evaluate, the experimental documentation may be incomplete or missing (especially in the case of experiments performed in the distant past), the procedures to achieve this are well established [Dean 2008], at least for the case of the ICSBEP [NEA ICSBEP] and IRPhE [NEA IRPhE] databases, and this information is usually provided in the benchmark documentation.

However, the uncertainties in measured quantities of interest (or "integral parameters") of different reactor systems are not necessarily independent, but they may be correlated. In fact, nuclear benchmark experiments are usually performed in series, either having been performed at the same experimental facility or at different facilities sharing some characteristics or properties, e.g. using the same materials, instrumentation, experimental techniques, etc. Consequently, the measured values of these experimental parameters will not be independent but will have some degree of correlation. Although these correlations are known to occur, relatively little attention has been paid to them, and no information about them is provided in ICSBEP benchmark documentation.

Having reached this point, it is important to distinghish between the concepts of similarity and correlation:

1) ***Similarity***. Some degree of similarity between experiments is due to two experiments being sensitive to the same set of nuclear data. For instance, two light-water moderated, uranium-fuelled experiments will likely be more similar between them than to a fast plutonium system. Hence, similarities are present between experiments which utilize the same types of materials, even if they are performed in totally independent facilities. Similarities are usually assessed via *similarity coefficients*. If a certain response of the two systems, k_1 and k_2, is a function of the same set of parameters $\alpha_1, \dots, \alpha_N$ (usually, neutron cross sections or other nuclear data) then the similarity in the response can be defined as:

$$E = \frac{S_{1,\alpha}^T S_{2,\alpha}}{\sqrt{(S_{1,\alpha}^T S_{1,\alpha})(S_{2,\alpha}^T S_{2,\alpha})}} = \frac{S_{1,\alpha}^T S_{2,\alpha}}{|S_{1,\alpha}||S_{2,\alpha}|} \qquad (2.1)$$

   Where $S_{1,\alpha} = (\partial k_1/\partial \alpha_1, \dots, \partial k_1/\partial \alpha_N)$ and $S_{2,\alpha} = (\partial k_2/\partial \alpha_1, \dots, \partial k_2/\partial \alpha_N)$ denote the sensitivity vectors of the responses k_1 and k_2 to the set of parameters $\alpha_1, \dots, \alpha_N$. Although these similarity coefficients play an important role in nuclear data adjustment and assimilation studies, they are out of the scope of this study.

2) ***Correlation in the experimental input parameters***. These are the correlations that are assessed in this work. They are not due to two experimental measurements being sensitive to the same parameters (nuclear data) but, as stated above, they arise when two experiments are performed in the same facility or otherwise share some commonality (e.g. the fuel composition in two experiments performed with the same fuel), the corresponding uncertainties will therefore not be independent but correlated.

3 Impact of experimental correlations

Experimental correlations have a major impact for at least two applications: (1) assessing the calculational bias in criticality safety assessment, and (2) nuclear data adjustment and assimilation. They are briefly reviewed in this section.

3.1 Calculational bias in Criticality Safety Assessment

The validation of neutron transport codes and nuclear data is an important step in criticality safety assessment. This validation is performed by comparing the calculated results of the multiplication factorwith the experimental results of benchmark experiments. In this context, if we have an integral experiment for which both experimental and calculated values of the multiplication factor k_{exp} and k_{calc} are available, the calculational bias can be defined as[1]:

$$\Delta k = k_{calc} - k_{exp} \quad (3.1)$$

If validation is performed using a set of fully independent integral experiments (i.e. assuming that they experimental parameters are fully uncorrelated, and hence their uncertainty in k_{exp}), then the average calculational bias and its variance will be given by the well-known expressions (see e.g. the Statistics section of [Tanabashi 2018]) :

$$\overline{\Delta k} = \frac{\sum_{i=1}^{N} \frac{\Delta k_i}{\sigma_{k_{exp,i}}^2}}{\sum_{i=1}^{N} \frac{1}{\sigma_{k_{exp,i}}^2}} = \frac{\sum_{i=1}^{N} \frac{k_{calc,i} - k_{exp,i}}{\sigma_{k_{exp,i}}^2}}{\sum_{i=1}^{N} \frac{1}{\sigma_{k_{exp,i}}^2}} \quad (3.2)$$

$$\sigma_{\overline{\Delta k}} = \frac{1}{\sqrt{\sum_{i=1}^{N} \frac{1}{\sigma_{k_{exp,i}}^2}}} \quad (3.3)$$

However, if some correlations exist between the experiments, these expressions will become [Schmelling 1995, Ivanova 2003a]:

$$\overline{\Delta k} = \frac{\sum_{i,j=1}^{N} C_{i,j}^{-1} (k_{calc,i} - k_{exp,i})}{\sum_{i,j=1}^{N} C_{i,j}^{-1}} \quad (3.4)$$

$$\sigma_{\overline{\Delta k}} = \frac{1}{\sqrt{\sum_{i,j=1}^{N} C_{i,j}^{-1}}} \quad (3.5)$$

Where $C_{ij} = \rho_{ij} \sigma_i \sigma_j$ is the $N \times N$ covariance matrix of the set of *N* integral experiments, and ρ_{ij} is the Pearson's correlation coefficient between the experiments *i* and *j*, defined as:

$$\rho_{ij} = \frac{cov(k_{exp,i}, k_{exp,j})}{\sqrt{var(k_{exp,i})} \sqrt{var(k_{exp,j})}} \quad (3.6)$$

As an example of the relevance of the correlations to determine the calcualtional bias, let us consider a simple case of two experiments with $k_{calc,1} - k_{exp,1} = 0.01$ and $k_{calc,2} - k_{exp,2} = 0.02$, with experimental

[1] In [NEA 2013a] eq. 92 an alternative definition is given as:

$$\Delta k = \frac{k_{calc}}{k_{exp}} - 1$$

uncertainties $\sigma_{k_{exp,1}} = 0.7\%$ and $\sigma_{k_{exp,2}} = 1\%$. With these conditions, if experiments 1 and 2 are not correlated, the correlation matrix and the calculation bias will be:

$$C = \begin{pmatrix} \sigma^2_{k_{exp,1}} & 0 \\ 0 & \sigma^2_{k_{exp,2}} \end{pmatrix} \text{ and } \overline{\Delta k} = 0.0133 \pm 0.0057$$

On the other hand, if experiments 1 and 2 share some components of the uncertainty so that the correlation coefficient is 0.85, then the correlation matrix and the calculation bias will be:

$$C = \begin{pmatrix} \sigma^2_{k_{exp,1}} & 0.85\sigma_{k_{exp,1}}\sigma_{k_{exp,2}} \\ 0.85\sigma_{k_{exp,2}}\sigma_{k_{exp,1}} & \sigma^2_{k_{exp,2}} \end{pmatrix} \text{ and } \overline{\Delta k} = 0.0065 \pm 0.0067$$

As a final comment, if a Monte Carlo code is used to determine the correlations, the calculated values will also be affected by statistical errors inherent to the Monte Carlo method (see section 5.2).

3.2 Nuclear data adjustment

Another field where the correlation matrices between experimental experiments are used is for nuclear data adjustment [NEA 2010, NEA 2013]. Nuclear data adjustment is a procedure to determine an improved value of a cross section from the results of integral experiments. In this procedure, if a set of N_σ cross section values are to be determined (N_σ = number of isotopes × number of energy groups) and a set of N_E experimental values, then the better values of the cross sections will be the ones minimizing the following expression:

$$\chi^2 = (\vec{\sigma} - \vec{\sigma}_m)^T M_\sigma^{-1} (\vec{\sigma} - \vec{\sigma}_m) + \left(\vec{E} - \vec{C}(\vec{\sigma})\right)^T M_E^{-1} \left(\vec{E} - \vec{C}(\vec{\sigma})\right) \quad (3.7)$$

Where

- $\vec{\sigma} = (\sigma_1, \dots, \sigma_{N_\sigma})$ is the adjusted (a posteriori) cross section vector.
- $\vec{\sigma}_m = (\sigma_{m,1}, \dots, \sigma_{m,N_\sigma})$ is the original (a priori) cross section vector.
- M_σ is the $N_\sigma \times N_\sigma$ covariance matrix of nuclear data.
- $\vec{E}$ is the N_E vector of experimental values.
- $\vec{C}(\vec{\sigma})$ is the N_E vector of calculated values, with the a-posteriori set of data $\vec{\sigma}$.
- M_E is the $N_E \times N_E$ experimental covariance matrix of the experimental values.

In addition, M_E can include not only the experimental uncertainties, but also the experimental and modelling covariances. For information about how to obtain the modelling uncertainties, see for instance [NEA 2013].

4 Status of existing correlations

4.1 Correlations between criticality benchmarks

Many benchmark experiments in the International Handbook of Evaluated Criticality Safety Benchmark Experiments (ICSBEP) and in the International Handbook of Evaluated Reactor Physics Experiments (IRPhE) provide detailed information about the uncertainties in the system description as well as their estimated impact on the integral responses. In the case of ICSBEP, this information is provided in section 2 *"Evaluation of experimental data"* of the benchmark description documents [Dean 2008].

To our knowledge, the most complete database of correlations between integral experiments is the one included in the NEA's DICE tool [NEA DICE, Nouri 2003]. However, even this one contains a rather limited amount of experimental correlation data: as of February 2024, correlation coefficient data between experimental benchmark criticality cases are available for only 93 cases in the DICE database associated with the ICSBEP handbook. This represents a small fraction of the total of over 5000 integral experiments, underscoring the fairly small number of evaluations that contain quantitative correlations.

This information can be extracted from DICE from the panel "Correlation Matrix → Display Uncertainties → Show cases level details". A qualitative correlation matrix is provided, where the degree of correlation goes from 0 and 1000, 0 meaning no correlation and 1000 meaning full correlation, i.e. Pearson's correlation coefficient is equal to 1 (Figure 1).

More specifically, the sets of integral experiments for which quantitative correlation data are present are:

- 55 cases corresponding to four sets of HEU-SOL-THERM benchmarks: 21 from IPPE (corresponding to the benchmarks HST019/025/027/028/029/030/035), 24 from ORNL, further divided into two blocks of 14 (HST009/ /010/011/012/043) and 10 (HST013/032/042) and 10 from Rocky Flats (HST001). The likely source is [Ivanova 2003a].
- A total of 33 cases from the Zero Power Reactor (ZPR) and/or Zero Power Physics Reactor (ZPPR) experiment series of Argonne National Laboratory. The likely source is [Palmiotti 2014].
- 3 cases from VNIIEF (HEU-MET-FAST-018-001, HEU-MET-FAST-020-001 and HEU-MET-FAST-031-001). They correspond to three experiments carried out at the CTF facility, a bare, 90% enriched uranium with different reflectors.
- 2 cases from VNIITF (HEU-MET-FAST-008-001, HEU-MET-FAST-011-001). They correspond to two different configurations from the FKBN facility: highly-enriched uranium spheres, both bare and with a polyethylene reflector.

A major limitation of DICE, in our opinion, is that no information about the source of these correlation data is provided (although in some cases it can be guessed, as stated above), and hence we do not know the methodology followed and the hypothesis made to determine them (i.e., which variables are considered to be correlated between these experiments, see section 5), which makes difficult to reproduce the results.

In addition to these experimental correlations included in DICE, quantitative results are available in the literature for some other systems:

- Correlations in k_{eff} of some selected fast experiments taken from ICSBEP and IRPhE: BFS-97/99/101 [Ivanova 2009] and Zebra, MZA, MZB, ZPR-6/7 and Sneak 7A/B [Ivanova 2014]
- Correlations for experiments involving arrays of low-enriched fuel rods measured at the "Apparatus-B" facility at CEA-Valduc (France), corresponding to the LEU-COMP-THERM-007 and LEU-COMP-THERM-

039 benchmarks of ICSBEP. These correlations were performed under EG UACSA Benchmark Phase IV and results calculated by different organizations are available [NEA 2023].

- Correlations between 7 cases of LEU-COMP-THERM-042, also consisting of arrays of water-moderated fuel rods, measured at the Critical Mass Laboratory of PNL at Hanford (USA). [Marshall 2017a, Marshall 2017c, Marshall 2019].
- Correlations for experiments involving high-enriched uranium solutions with a thermal neutron energy spectrum. 1 evaluation was considered: HEU-SOL-THERM-001 [Marshall 2017b, Marshall 2017c].
- Correlations for plutonium nitrate in aqueous solution contained in metal sphere. Analysis of 43 critical experiments from 6 experimental series PU-SOL-THERM-03 to -06, and -20 and -21 [Kilger 2016].
- Correlations between the keff of the ZPR-6/7, ZPR-6/7 with high Pu-240 content and ZPPR-9 experiments of ANL (USA) [NEA 2013b]. This reference also list the correlation between other experimental parameters of fast reactors (spectral indexes, reactivity effects).

A summary list of these references, classified by the methodology used to determine them (deterministic and Monte Carlo), is given in Table 1. It is worth stressing here that different evaluations of the experimental correlations can produce very different results. For instance, the correlation between the k_{eff} of ZPR-6/7 (MCF001-001) and ZPR-6/7 with high Pu-240 content (MCF002-001) is listed as 0.66 in DICE and [Palmiotti 2014], but as 0.134 in [Ivanova 2014] and [NEA 2013b].

Finally, in addition to the systems for which DICE provides quantitative correlations listed above, qualitative correlations are provided for a larger set of cases (Figure 2). A symbol ‘‘+” indicates that these two cases are correlated or, in other words, that one or several uncertain benchmark parameters that are major contributors to the overall experimental k_{eff} uncertainty are correlated. The symbol ‘‘(+)” indicates a 100% correlation, hence it usually appears in the diagonal. When it appears outside the diagonal, it means that an evaluation has more than one identifier in ICSBEP, i.e. is a cross-reference. Some of the systems for which the existence of correlations is stated in DICE – and are therefore worth investigating – are:

1) IPPE: 30 benchmark experiments in three facilities (BRR, KBR and BFS). Among them, correlations between 10 configurations of BFS are provided in [Ivanova 2009], as stated above.
2) VNIIEF: 34 benchmark experiments performed at the CTF facility, which can be further classed in 6 groups according to the fuel used. In DICE, quantitative correlations only for three cases are listed.
3) VNIITF: 54 benchmarks experiments performed at the FKBN facility, also classified in four groups according to the fuel used and the geometry (cylindrical or spherical). Quantitative correlations are only listed in DICE for four cases.
4) Kurchatkov Institute: 30 benchmark in 6 groups, according to the experimental facility.
5) JAEA: 29 benchmark experiments in two facilities (5 in TCA and 24 in STACY)
6) IPEN: 18 benchmark experiments in the MB-01 reactor.

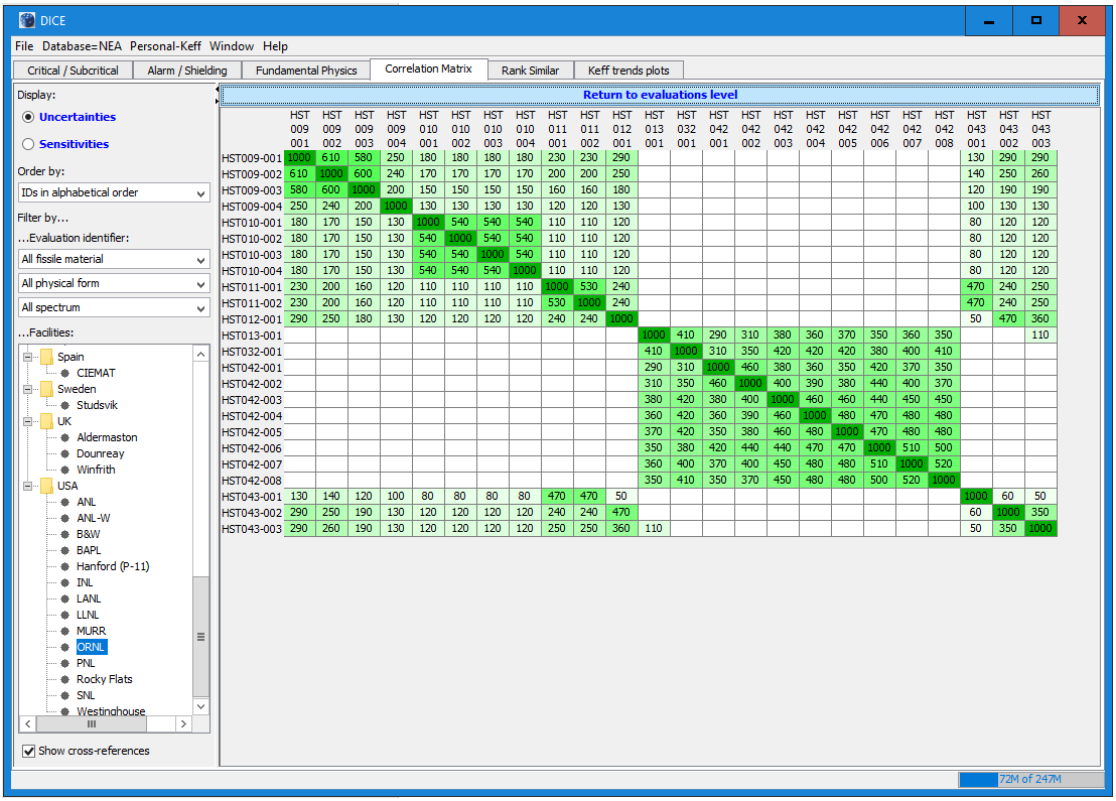


Figure 1. Example of quantitative correlations between experimental benchmarks provided by DICE.

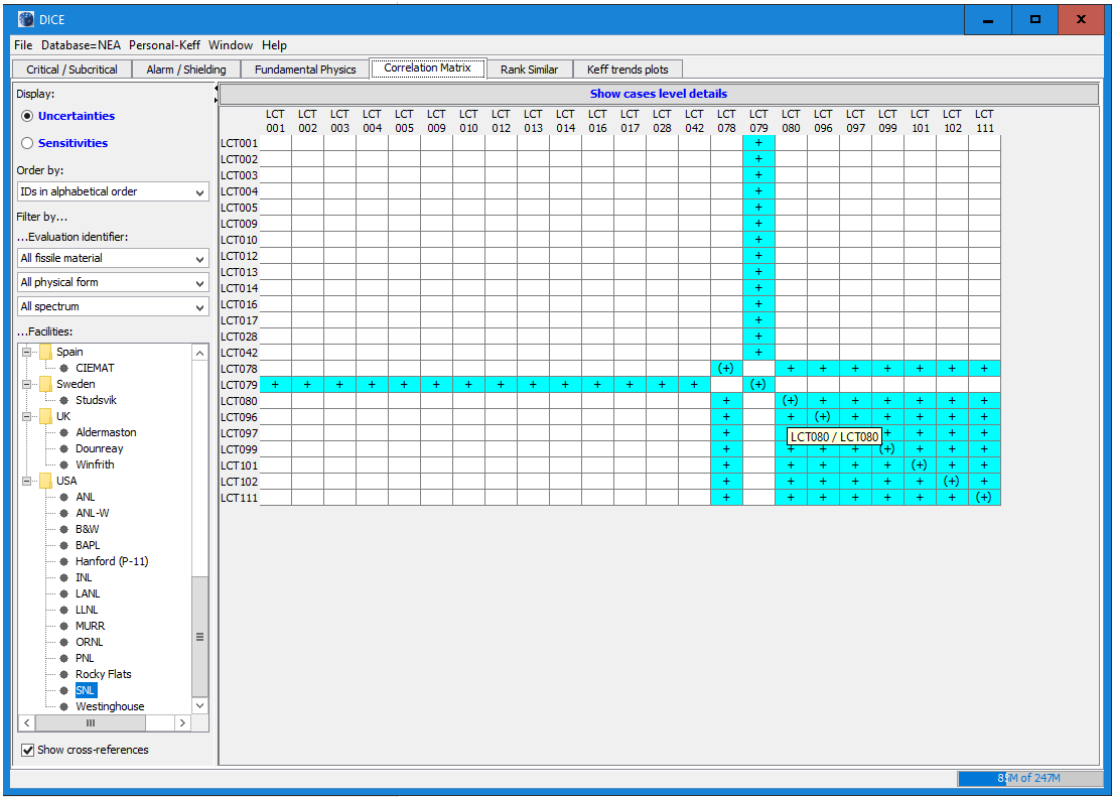


Figure 2. Example of qualitative correlations between experimental benchmarks provided by DICE.

4.2 Correlations between shielding benchmarks

Although less attention has been paid to them, correlations between parameters in different shielding experiments are of interest as they involve other types of integral measurements. In fact, one of the strong points of the SINBAD shielding benchmark project [NEA SINBAD, Kodeli 2013] is that it devotes great attention to the careful evaluations, and the understanding and consolidation of the uncertainties involved in the measurements, stressing the importance of different terms and types of uncertainties. Where the information is available, the systematic and stochastic components of the uncertainties are differentiated and analysed, which allows to establish the full covariance matrix among the measured values. Among the good examples where this was possible, the ASPIS, FNG and others benchmarks can be cited. As an example, the covariance matrix of the reaction rates measured in the ASPIS Iron-88 benchmark [Wright 1993, Avery 1995, Milocco, 2015] is presented in section 6.3.

In section 6.3, it is demonstrated another useful approach: in some cases, if the information on the systematic uncertainty is available (and reliable), it is advantageous to compare the ratios of the measured quantities, as the (major part of) the systematic uncertainty then cancels out. It is worth remarking, however that the statistical uncertainties, both experimental and computational (e.g. Monte Carlo statistics) are summed up. This approach, although often used, may also lead to the loss of some experimental information (e.g. on the spectra distribution, which is due to the use of the “case-dependent” normalisation factor).

5 Methodologies to generate correlations

5.1 Deterministic approach

The deterministic methodology was first proposed in [Ivanova 2003]. Let us consider a set of N criticality benchmark experiments whose multiplication factor *k* depends on a set of M experimental parameters (e.g. compositions, dimensions, etc.). As stated above, these parameters are inevitably affected by experimental uncertainties, which, in turn, will propagate into an uncertainty in *k*. Let us denote by $\delta k_{\alpha,i}$ the uncertainty in the multiplication factor k_i of the i-th experimental benchmark due to the uncertainty in the experimental parameter α. If, for this i-th integral experiment, we can assume that the set of parameters α are uncorrelated[2] the total uncertainty in k_i will be given by:

$$\delta k_i = \sqrt{\sum_\alpha \delta k_{\alpha,i}^2} \qquad (5.1)$$

Then, if a parameter α is correlated between the experiments *i* and *j*, the correlation coefficient between the uncertainties of the multiplication factor of two experiments, k_i and k_j, in this set of benchmarks can be calculated as:

$$\rho_{ij} = \frac{\sum_\alpha \delta k_{\alpha,i} \gamma_\alpha^{i,j} \delta k_{\alpha,j}}{\delta k_i \delta k_j} = \frac{\sum_\alpha \delta k_{\alpha,i} \gamma_\alpha^{i,j} \delta k_{\alpha,j}}{\sqrt{\sum_\alpha \delta k_{\alpha,i}^2}\sqrt{\sum_\alpha \delta k_{\alpha,j}^2}} \qquad (5.2)$$

Where $\gamma_\alpha^{i,j}$ denotes the correlation between the uncertainties in the parameter α between the experiments *i* and *j*, i.e. $\gamma_\alpha^{i,j} = 1$ if the uncertainties are fully correlated (e.g. the uncertainty in the fuel composition in two experiments sharing the same fuel) and $\gamma_\alpha^{i,j} = 0$ if the uncertainties are fully independent (e.g. uncertainties in the fuel composition in two experiments using different fuels). If $i = j$, then $\gamma_\alpha^{i,i} = \gamma_\alpha^{j,j} = 1$ and $\rho_{ii} = \rho_{jj} = 1$. Notice, however that if $i \neq j$, the values ρ_{ij} will be, in general, lower than 1 even if all parameters α are fully correlated[3]. We use the notational convention that Greek letters range over the experimental parameters and Latin letters over experiments.

The values $\delta k_{\alpha,i}$ can be calculated with neutron transport codes. In fact, they are often provided in the benchmark documentation (section 2 of ICSBEP benchmarks, *"Evaluation of experimental data"*). Having reached this point, it is worth mentioning that Eq. 5.2 can be further developed by expressing the uncertainties $\delta k_{\alpha,i}$ in terms of sensitivities of k_i to the parameter α and the uncertainty in α itself ($\delta\alpha$), namely:

$$\delta k_{\alpha,i} = \left(\frac{\partial k_{\alpha,i}}{\partial \alpha}\right)\delta\alpha = S_{\alpha,i}\delta\alpha \qquad (5.3)$$

[2] In principle it is possible to achieve a description of a system in terms of uncorrelated parameters if only the "fundamental" parameters of the system are considered. But in practice, many times some "intermediate" parameters, that may already include correlations, are taken as inputs in some simplified model.

[3] This arises from the fact that different benchmark experiments will have in principle different sensitivities to a given parameter. Consider, for instance two benchmark systems sharing two experimental parameters so that the uncertainties in k in the two systems due to these parameters are $\delta k_{\alpha=1,i=1} = 0.05$, $\delta k_{\alpha=2,i=1} = 0.2$, $\delta k_{\alpha=1,i=2} = 0.1$, $\delta k_{\alpha=2,i=2} = 0.1$. Then,

$$\rho_{12} = \frac{\delta k_{\alpha=1,i=1}\delta k_{\alpha=1,i=2} + \delta k_{\alpha=2,i=1}\delta k_{\alpha=2,i=2}}{\sqrt{\delta k_{\alpha=1,i=1}^2 + \delta k_{\alpha=2,i=1}^2}\sqrt{\delta k_{\alpha=1,j=1}^2 + \delta k_{\alpha=2,j=1}^2}} = 0.8575$$

Where the experimental uncertainty $\delta\alpha$ has to be obtained from the experiment documentation and the sensitivities $S_{\alpha,i}$ can be obtained with neutron transport codes[4]. With this notation, Eq. 5.2 can be expressed in the matrix form:

$$\rho_{ij} = \frac{\sum_\alpha S_{\alpha,i}\delta\alpha\gamma_\alpha^{i,j}\delta\alpha\delta S_{\alpha,i}}{\delta k_i \delta k_j} = \frac{(S_{1,i} \quad \dots \quad S_{M,i})\begin{pmatrix}\gamma_1^{i,j}\delta\alpha_1^2 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \gamma_M^{i,j}\delta\alpha_M^2\end{pmatrix}\begin{pmatrix}S_{1,j}\\ \vdots\\ S_{M,j}\end{pmatrix}}{\delta k_i \delta k_j} = \frac{\vec{S}_i^T V^{ij}\vec{S}_j}{\delta k_i \delta k_j} \qquad (5.4)$$

V^{ij} being the $M \times M$ covariance matrix of the set of experimental parameters. Notice that the terms outside the diagonal are zero because we have considered that the parameters α are not correlated for a given experiment; should correlations within a given experiment exist, the corresponding non-diagonal terms will be non-zero. This matrix formulation is known as the "sandwich rule" (e.g. [Cacuci 2003]) and is used in e.g. [Dos Santos 2013]. Notice that in [Dos Santos 2013], a further distinction is made between "technological uncertainties" and "measurement technique uncertainties" resulting in a formula of the type:

$$\rho_{ij} = \frac{\vec{S}_i^T V_{tech}^{ij}\vec{S}_j + \delta_{exp}^{ij}}{\delta k_i \delta k_j} \qquad (5.5)$$

Where the measurement technique correlation matrix takes the form of the Kronecker symbol: it was considered to be 1 if the technique used to measure the multiplication factor *k* is the same in the two techniques, or 0 if two different techniques were used.

As stated above, this technique was first applied, to our knowledge, in [Ivanova 2003]. A summary of other works applying this deterministic methodology to find correlations between experimental benchmarks is presented in Table 1(a).

[4] These sensitivity coefficients can be determined by comparing 2 direct calculations or by using perturbation theories. Many modern neutron transport codes can perform perturbation calculations (e.g. KPERT and KSEN functionalities in the MCNP code), thus allowing to carry out several such calculations.

Table 1. Correlations between integral experiments available in the literature.

(a) Obtained with the deterministic methodology.

References	Summary
[Ivanova 2003a] [Ivanova 2003b]	Methodology described. Correlations in k_{eff} unc. for 77 HEU-SOL-THERM cases: 34 from IPPE, 10 from Rocky Flats, 29 from ORNL and 4 from LANL. No numerical information.
[Ivanova 2009]	Correlations in k_{eff} unc. for 10 cases from IPPE's BFS-99, 99 and 101 (MMCF003/004, MMCM001)
[Dos Santos 2013] [Dos Santos 2014]	Methodology described. Correlations in k_{eff} unc. for 6 systems, including 3 cases from ICSBEP (ZPR-10A(?) and MCT001).
[Ivanova 2014]	Correlations in k_{eff} unc. for some fast benchmarks from IRPhE (ZEBRA, ZPR, SNEAK, NEA-NSC-WPEC-SG33) and thermal benchmarks for ICSBEP (LCT007-039, UACSA Benchmark Phase IV)
[NEA 2013b] [Salvatores 2014]	Correlations in k_{eff} unc. for ZPR-6/7 and ZPPR-9, correlations in spectral indexes in some other systems.
[Palmiotti 2014]	Correlations in k_{eff} unc. for 33 ZPR benchmark experiments. Results of US DOE Nuclear Data Adjustment Project.
[Jeong 2017]	Correlation matrices for some LEU-COMP-THERM and HEU-COMP-FAST cases. No numerical information.

(b) Obtained with the Monte Carlo methodology.

References	Summary
[Buss 2010]	Correlations in k_{eff} unc. for 97 LEU-COMP-THERM and MIX-COMP-THERM cases. MC code used: SCALE 6. No numerical information.
[GRS 2016b] [Peters 2016] [Kilger 2016]	Correlations in k_{eff} unc. for 9 LCT cases (LCT006 and LCT035/062, JAEA's TCA) and 43 PST cases (PST003/006/020/21) (no numerical information). MC code used: KENO.
[Marshall 2015] [Marshall 2017a] [Marshall 2017b] [Marshall 2017c]	Correlations in k_{eff} unc. for a series of cases of LCT007/039 (CEA Valduc), LCT042 (PNL) and HST001 (Rocky Flats) benchmarks. MC code used: KENO
[Sommer 2021]	S2Cor methodology for efficient MC sampling to calculate correlatons. Applied to some LCT007 (CEA Valduc) cases. MC code used: KENO

(c) Other sources of correlations between integral experiments.

References	Summary
[GRS 2016a] [Marshall 2017a] [Marshall 2017c] [Marshall 2019] [Stuke 2019] [NEA 2023]	Results of EG UACSA Benchmark Phase IV. Intercomparison of methodologies for generating correlation matrices in LCT-007 (4 cases) and LCT-039 (17 cases) (Apparatus-B @ CEA Valduc)

The most critical point in this approach is determining the correlation coefficients $\gamma_{\alpha}^{i,j}$. This is usually achieved by expert-judgement, since a high level of understanding of the experimental procedures is required to estimate the correlation degree. A systematic approach has been proposed in [Salvatores 2013], consisting of the following steps:

1) Classification of Error Components to either Common or Independent
2) Summation of Common and Independent Errors
3) Evaluation of correlation factor

5.2 Monte Carlo sampling approach

Methodologies based on random sampling of perturbed parameters within model inputs can be used to evaluate correlations. After determining the most important parameters influencing the uncertainty in the quantities of interest of the systems being investigated, and deciding which of them are correlated between the systems and which of them are not, these parameters are randomly sampled, taking into account the correlations. Then, calculations are performed for every system and set of parameters. Notice that the neutronic calculations are not necessarily performed with MC codes, but deterministic codes can be used as well. Then, eq. 3.6 is applied to the results of these calculations to obtain the correlation coefficients between every pair of systems.

The advantages of the Monte Carlo technique over the deterministic technique are that it does not require to know beforehand, calculate or make hypothesis about the values $\delta k_{\alpha,i}$. However, the issue of determining which ones are correlated or not remains open. The major disadvantage, as usual with any stochastic procedure, is the computational resources required. Related with this, important attention has to be paid to the convergence of correlation coefficients and the impact of stochastic uncertainty in the results. In appendix 1, equations to propagate these uncertainties are derived.

To our knowledge, the Monte Carlo technique was first applied to obtain correlation coefficients between experimental uncertainties nuclear benchmark experiments in [Buss 2010], using the SCALE code system. Other applications of the technique for finding experimental correlations are summarized in Table 1(b).

5.3 Machine Learning techniques for interpretation of experimental measurements

Recently, Machine Learning (ML) techniques have arisen as powerful methods capable to exploit the known information about some data sets (referred as training sets) to predict properties for an another, unknown data sets. The relationship between predicted properties and augmented variables is deduced from the training set (e.g. measured set of data) and is then applied to the non-training set (missing data points). The model is typically constructed by a Gaussian process, defined by mean values and covariance functions. The main challenge of using ML methods is the understanding and physical interpretation of the results obtained, involving both the original uncertainties (e.g. experimental) and ML model uncertainties.

As an example of the possible use and performances of different ML methods, an example is presented in section 6.4, where these techniques have been applied to exploit the experimental power distribution results measured at several location in the VENUS experimental reactor core during the VENUS-3 shielding benchmark [Berger 2024].

6 Applications

In this section, we present three applications of the previously described methodologies to determine the correlation coefficients between integral experiments: (1) correlations between two different configurations of the core of the EOLE light water zero-power reactor at CEA-Cadarache; (2) correlations between a set of experiments performed at the ZPR/ZPPR reactor at Argonne National Laboratory that have been obtained from the ICSBEP database, and (3) covariances between reaction rates measured at the ASPIS Iron-88 benchmark experiment. Furthermore, a recent example of the the application of Machine Learning techniques in the interpretation of uncertainties and correlations in a benchmark experiment (VENUS-3 shielding benchmark) is also presented.

6.1 EOLE/CAMELEON

The CAMELEON core was loaded in the EOLE reactor vessel (95cm of diameter) at the beginning of the 1980s [Santamarina 1982]. It was made of 805 UOx fuel pins (labelled UOX100 in Figure 3) for the central part and made of 959 UOx fuel pins (labelled UOX140 in Figure 3) for the peripheral part. All fuel rods, 3.5% ^{235}U enriched, were arranged in a square lattice (pitch was 1.26cm). The homogeneous loading plan, labelled "Tout UO_2", is shown in Figure 3. The critical boron was adjusted for each configuration.

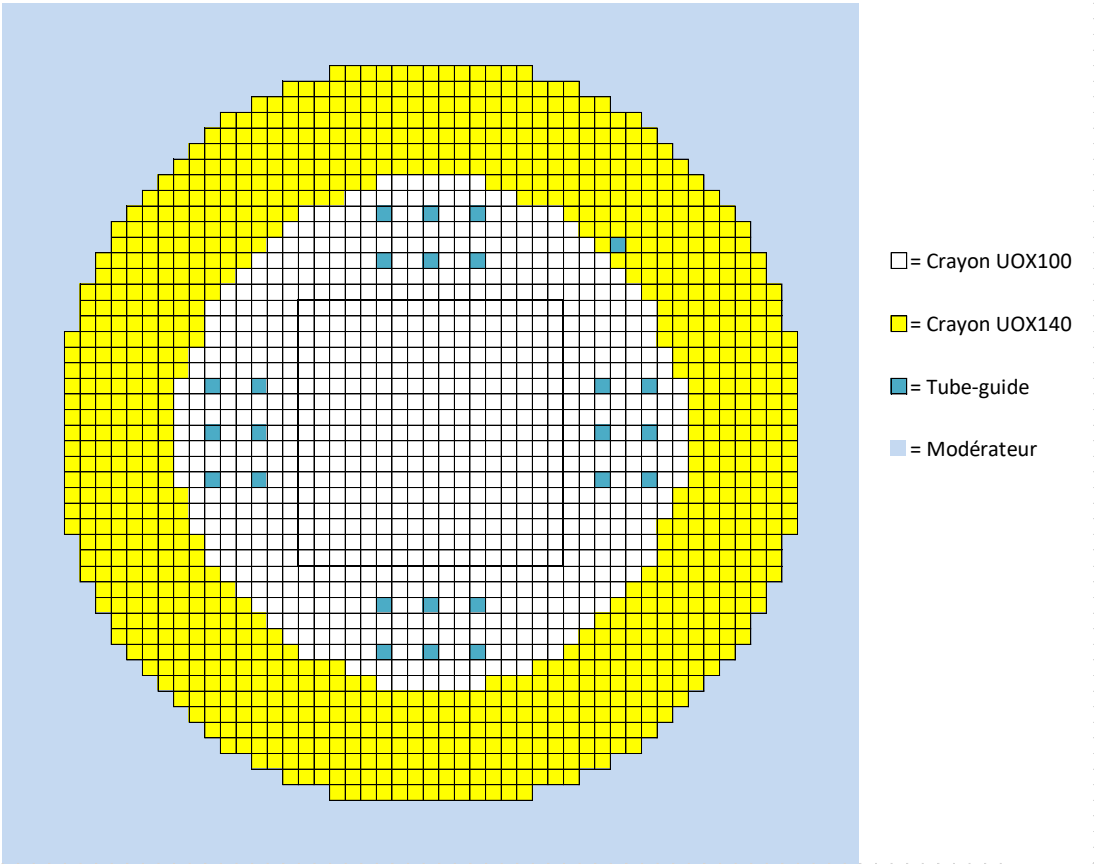


Figure 3. Reference configuration "Tout UO_2" of the CAMELEON core.

In addition to this homogeneous configuration, two other configurations with various contents of gadolinium in the center fuel were loaded into the EOLE core. As shown in Figure 4, the central 17x17 "assembly" was loaded with different patterns containing different number of gadolinium rods of two types:

- Type A, depleted uranium ^{235}U (0.25%) and 7% of Gd_2O_3,
- Type B, enriched ^{235}U (5%) and 3% of Gd_2O_3.

In this study, we decided to use the configuration of type A in order to maximize the gadolinium content. All fission rates are normalized to the sum of the 17 fuel pins fission rates labelled by '•' in Figure 4. Table 2 lists the measured integral data and associated experimental technique for the 3 different configurations considered here.

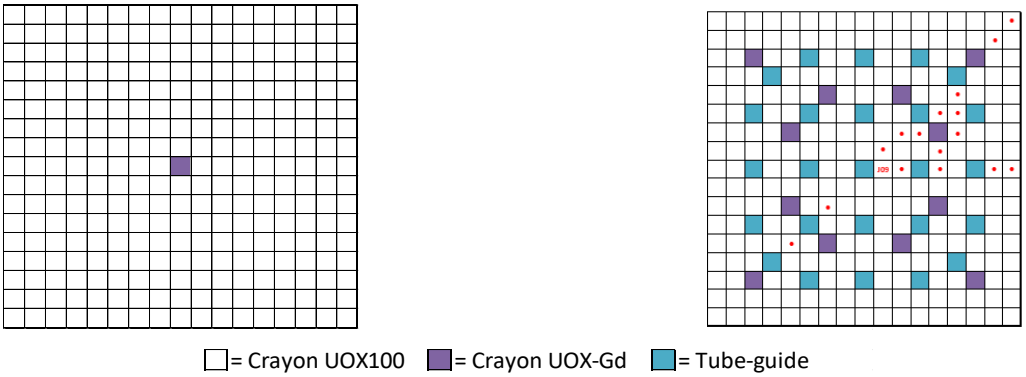


Figure 4. Central assembly of CAMELEON : "1Gd" rod and "12D3" 12 gadolinium rods. The J09 fuel pin used for this study is located in front of the central water hole.

Table 2. Studied configurations and observables.

Configuration	Integral measurement	Experimental technique
Tout UO_2	Residual reactivity	Doubling time
1-Gd-A	Reactivity effect (versus Tout UO_2 config)	Doubling time difference
12D3-A	Residual reactivity Radial fission rate	Doubling time γ-spectrometry

In the next section, we discuss the technological parameters and their associated covariances to be propagated to the core level.

6.1.1 Technological parameter identification and uncertainty propagation

The list of selected technological parameters that influence integral data is presented in Table 3. All these parameters are supposed to be independent (i.e., the covariance matrix is diagonal). The propagated uncertainties onto integral observables are then obtained by performing a quadratic sum of direct parameter perturbations using the sensitivity vectors calculated by the APOLLO3® neutronic code and a 2D spatial description. The resulting integral uncertainties are given in Table 4.

Most influent parameters are listed in bold: ^{235}U enrichment, external radius of fuel and cladding, lattice pitch and UO_2 impurities. UO_2+Gd technological parameters have a low propagated uncertainty on integral quantities because of their low number (only 12 Gd pins versus a total of 1700 UO_2 fuel pins).

Not mentioned in Table 4, the technological uncertainty for a reactivity effect due to a small modification (from "Tout UO2" configuration to a gadolinated configuration) is vanishing to zero.

Table 3. The list of selected technological parameters, and 1σ associated uncertainties.

CAMELEON technological parameters		1σ uncertainty (absolute / *relative*)
UOx fuel pin type 100	Isotopic enrichment ^{235}U	**0.02 [w/o unit]**
	Fuel density	**0.0075 [w/o unit]**
	Oxygen Stoichiometry in UO_2	*0.20%*
	Fuel impurities	**1 ppm of B [(5)]**
	Fuel radius	**0.01 [cm]**
	Cladding external radius	**0.01 [cm]**
UOx fuel pin type 140	Isotopic enrichment ^{235}U	**0.002 [w/o unit]**
	Fuel density	**0.0075 [w/o unit]**
	Oxygen Stoichiometry in UO_2	*0.20%*
	Fuel impurities	**1 ppm of B [(1)]**
	Fuel radius	**0.03 [(6)] [cm]**
	Cladding external radius	**0.01 [cm]**
UOx+Gd 7% fuel pin (type A)	Isotopic enrichment ^{235}U	**0.0005 [w/o unit]**
	Fuel density	**0.0075 [(7)] [w/o unit]**
	Oxygen Stoichiometry in UO_2	*0.20%*
	Fuel impurities	*0.10%*
	Fuel radius	*0.20%*
	Ratio ^{160}Gd / ^{238}U	**12.5 [ppm]**
General parameters	Lattice pitch	**0.002 [cm]**
	Water density	*0.01%*

(5) The level of impurities in the fuel is neutronically equivalent to a few ppm of boron.
(6) This value is obtained by observing the measurements standard deviation.
(7) This UOx value is arbitrary given to UOx+Gd fuel pins.

Table 4. k_{eff} and fission rate technological uncertainties.

CAMELEON technological parameters		Technological uncertainty	
		k_{eff} (12D3-A) [pcm]	Fission rates (12D3-A) min / J09 / max [%]
UOx fuel pin type 100	Isotopic enrichment ^{235}U	**80**	**0.22 / 0.31 / 0.34**
	Fuel density	4	0.02 / 0.03 / 0.04
	Oxygen Stoichiometry in UO_2	2	0.00 / 0.01 / 0.01
	Fuel impurities	**62**	**0.13 / 0.16 / 0.16**
	Fuel radius	**44**	**0.20 / 0.31 / 0.35**
	Cladding external radius	**53**	**0.30 / 0.35 / 0.42**
UOx fuel pin type 1040	Isotopic enrichment ^{235}U	4	0.01 / 0.01 / 0.01
	Fuel density	4	0.00 / 0.01 / 0.01
	Oxygen Stoichiometry in UO_2	0	0.00 / 0.00 / 0.00
	Fuel impurities	**33**	0.05 / 0.09 / 0.09
	Fuel radius	**66**	0.01 / 0.06 / 0.06
	Cladding external radius	**31**	0.00 / 0.01 / 0.02
UOx+Gd 7% fuel pin (type A)	Isotopic enrichment ^{235}U	0	0.00 / 0.00 / 0.00
	Fuel density	1	0.00 / 0.00 / 0.00
	Oxygen Stoichiometry in UO_2	0	0.00 / 0.00 / 0.00
	Fuel impurities	4	0.00 / 0.00 / 0.00
	Fuel radius	0	0.01 / 0.02 / 0.04
	Ratio ^{160}Gd / ^{238}U	0	0.00 / 0.00 / 0.00
General parameters	Lattice pitch	**83**	**0.37 / 0.37 / 0.50**
	Water density	2	0.01 / 0.01 / 0.01
Total uncertainty		173 pcm	0.69% / 0.70% / 0.73%

6.1.2 Correlation between integral data

Technological correlation matrix of 2 experimental data for the CAMELEON program is obtained by using equation 5.5 and is presented in Table 5.

Table 5. CAMELEON experimental correlation coefficients.

	k_{eff} (12D3-A)	Δρ (1-Gd-A)	**J09** fission rate (12D3-A)
k_{eff} (12D3-A)	1.000	0.637	0.445
Δρ (1-Gd-A)		1.000	0.000
J09 fission rate (12D3-A)			1.000

The physical meaning of this matrix is interesting: residual reactivity measurement is not that much correlated to the measurement of gadolinium worth (0.4), meaning that these 2 observables depend differently on technological parameters (small collinearity of sensitivity vectors).

6.2 Experiments useful for sodium fast reactors

Validation of nuclear data for advanced sodium fast reactors requires assessing the library performance for fast spectrum experimental benchmarks. Taking as reference systems both the ESFR core design (European Sodium Fast Reactor) and ASTRID core design, an analysis of ICSBEP and IRPhE databases has been performed to find experiments useful for validation [García-Herranz 2024]. As a result of this analysis, a set of 34 integral experiments of ICSBEP have been identified as able to provide relevant information for SFR. They correspond to benchmarks with plutonium (PU) and mixed plutonium-uranium (MIX) fissile materials, with a physical form of compound (COMP) or metal/alloy (MET) and with a FAST neutron spectrum. The following correlations have been found for these cases:

- Correlation between MCF001 and MCF002 is provided in DICE (0.66). However, no quantitative correlation is given with other ANL evaluations (MCF005 and MCF006), although the existence of a qualitative correlation is indicated (+). Furthermore, as stated before, DICE's correlation value seem to be taken from [Palmiotti 2014]. However, in [Ivanova 2014] and [NEA 2013b] an entirely different value for the same correlation (0.13) is provided.
- Correlation between MCF001 and PMF033-001 is provided; however, neither quantitative correlation nor qualitative is given with PMF-001, -002, -005, -006, -008, -044 in LANL; with PMF-003 and -017 in LLNL; with PMF-037 at CML in Rocky Flats; with PMF-021, -025, -029, -030, -032, -035 cases in VNIIEF, Russia.
- Correlations between MCF-001 and -002 with MMF-011-001 and -002, -003 and -004 are included.

Considering the previous information, an analysis of the MCF experimental benchmarks carried out at ANL (from MCF001 to MCF006) has been performed. In the cases where correlations are available in DICE, the obtained correlation coefficients will be compared to them.

These experiments were performed at the ZPR facility of Argonne National Laboratory (ANL) between 1966 and 1971. The facility consisted of a large matrix made of steel that could be filled with drawers containing different materials, either nuclear fuel or other materials to simulate coolants, reflectors or other structural components of fast reactors. More specifically, the matrix was made of two parts, one fixed and the other movable, and criticality was achieved by joining both.

For this, we have used the Monte Carlo methodology described in section 5.2. We have first carried out an analysis of the benchmark documentation to select the correlated experimental parameters with a significant impact on the uncertainty in k_{eff} (section 6.2.1). Once these parameters are identified, random perturbations of these parameters are performed in the input files, taking into account the correlation (i.e., if a certain parameter is considered to be correlated between two different experiments, the same perturbation is performed in the two experiments, if not, it is perturbed independently). The neutronic calculations have been performed with the MCNP6.2 code [LANL 2017] and the JEFF-3.3 library [Plompen 2020] processed with the NJOY-21 code [NJOY 2021]. More details about the calculation procedure and the results obtained are provided in section 6.2.2.

One difficulty that has been found during the calculations is that the MCNP inputs provided with ICSBEP are homogenized in a reduced number of regions, e.g. core (inner and outer), blanket, reflector… This makes it an issue to find the correlations between two experiments having some shared components in one of these regions and onot shared in another region, e.g. two different types of fuel. In these cases, a guess is made about which is the dominant component, and the parameter will be considered correlated or not accordingly.

Some characteristics of the selected experiments are summarized in Table 6. Overall, it can be observed that two main types of Pu fuel, from two different manufacturers (Sefor and Dow/Numec) were used in the core:

- A mixture of Pu-U-Mo fuel from Sefor (approx. 91% ^{239}Pu and 9% ^{240}Pu) and Pu-Al fuel (95% ^{239}Pu, 5% ^{240}Pu). Used in the MCF003-001/002 experiments, carried out in 1966/67.
- Pu-U-Mo fuel from Dow and Numec, with approx. 87% ^{239}Pu and 11% ^{240}Pu. Used in all remaining experiments, all carried out after 1968. In MCF-004-001 only fuel from Dow was used, but in our calculations, we have made no distinctions between this case and the others. Furthermore, some Pu-Al fuel was also used in the core periphery of MCF001-001, but it is expected to have a small impact on the reactivity.

Furthermore, the MCF002-001 and MCF003-002 experiments were partly loaded (innermost part of the cores) with fuels with higher ^{240}Pu content.

Regarding the uranium fuel, all the uranium present in these systems was in the form of depleted uranium. However, it was present in three different chemical forms: as part of the Pu-U-Mo fuel, as U_{MET} and U_3O_8.

Table 6. Main characteristics of the MIX-COMP-FAST benchmarks considered in this study.

Benchmark	Year	Description
MCF001-001 ZPR-6/7 loading 12	1970	Core: Pu-U-Mo (87% ^{239}Pu, 11% ^{240}Pu, 0.2% ^{235}U), depleted U_3O_8, Na, Fe_2O_3 Core boundary: Pu-Al (95% ^{239}Pu, 5% ^{240}Pu), ... Reflector: depleted U_3O_8
MCF002-001 ZPR-6/7 loading 99 (high ^{240}Pu)	1970-1971	Inner core: Pu-U-Mo (69% ^{239}Pu, 26% ^{240}Pu, 0.2% ^{235}U), depleted U_3O_8, Na, Fe_2O_3 Outer core: Pu-U-Mo (87% ^{239}Pu, 11% ^{240}Pu, 0.2% ^{235}U), depleted U_3O_8, Na, Fe_2O_3 Core boundary: Pu-Al (95% ^{239}Pu, 5% ^{240}Pu) Reflector: depleted U_3O_8
MCF003-001 ZPR-3/48	1966	Core: Pu-U-Mo (91% ^{239}Pu, 9% ^{240}Pu, 0.2% ^{235}U), Pu-Al (95% ^{239}Pu, 5% ^{240}Pu), depleted U_{MET}, Na, graphite Blanket: depleted U_{MET}, Al
MCF003-002 ZPR-3/48B	1967	Inner core: Pu-U-Mo (91% ^{239}Pu, 9% ^{240}Pu, 0.2% ^{235}U), Pu-Al (75% ^{239}Pu, 22% ^{240}Pu), depleted U_{MET}, Na, graphite Outer core: Pu-U-Mo (91% ^{239}Pu, 9% ^{240}Pu, 0.2% ^{235}U), Pu-Al (95% ^{239}Pu, 5% ^{240}Pu), depleted U_{MET}, Na, graphite Blanket: depleted U_{MET}
MCF004-001 ZPR-3/56B	1968-1969	Core: Pu-U-Mo (87% ^{239}Pu, 11% ^{240}Pu, 0.2% ^{235}U), depleted U_3O_8, Na, Na_2CO_3, Fe_2O_3 Reflector: Ni, Na, Al
MCF005-001 ZPR-9/31	1976	Core: Pu-U-Mo (87% ^{239}Pu, 11% ^{240}Pu, 0.2% ^{235}U), depleted U_{MET}, Na, graphite Blanket: depleted U_{MET}, Na, graphite Reflector: steel
MCF006-001 ZPPR-2	1970-1971	Core: Pu-U-Mo (87% ^{239}Pu, 11% ^{240}Pu, 0.2% ^{235}U), depleted U_3O_8, Na, Fe_2O_3 Blanket: depleted U_3O_8, depleted U_{MET}, Na, Na_2CO_3, Fe_2O_3, steel (304) Reflector: steel

6.2.1 Correlated parameter identification

To determine which parameters have the highest impact on the uncertainty in k_{eff}, we have used the uncertainty information provided in section 2 of the benchmark description. We have made no dedicated calculations for this purpose. From this information, we have been able to determine four major components that jointly are responsible for about 90% of the uncertainty in k_{eff} or higher: the mass of Pu fuel, the mass of U, the ^{235}U content of U, the matrix tube pitch and the mass of iron in the matrix. The values are summarized in Table 7.

In the benchmark documentation, a very detailed description is included about the different components of the uncertainty in keff. For this work, we have made some simplifying assumptions:

1) The uncertainty in keff due to Pu is mainly due to the uncertainty in Pu mass, which is stated to be 0.15%. This uncertainty is considered to be fully systematic and hence fully correlated between experiments sharing the same type of fuel. The uncertainty in the Pu composition has not been considered. Four types of Pu fuel have been considered: Pu from Sefor, Pu from Dow/Numec and the two fuels with higher 240Pu content that were loaded in the inner cores of MCF002-001 and MCF003-002.
2) The dominant sources of uncertainty in the uranium composition are the total mass of uranium (0.15%) and the uncertainty in the amount of ^{235}U (0.22% ± 0.01%). These apply to the U in the Pu-U-Mo plates, the U_3O_8 plates and the U_{MET} plates. In any case, the impact of the Pu-U-Mo plates on k_{eff} is the smallest of these three materials, so the uncertainty in the uranium composition is considered correlated for systems containing either U_3O_8 plates or the U_{MET} plates, and uncorrelated between the two systems containing U_3O_8 and U_{MET}. Only MCF006-001 contain both, but since the impact of U_3O_8 seems to be dominant, is considered as having only U_3O_8.
3) The major contributor to the uncertainty due to the iron & steel components comes from the mass of the matrix tubes, which are assumed to be common to all experiments and therefore to be fully correlated. The uncertainty in the iron mass of the of the matrix is stated to be 2%. Furthermore, for simplicity, only the composition of iron isotopes is changed.
4) The matrix tube pitch is 2.175 ± 0.001 in. in the X-direction and 2.277 ± 0.001 in. in the Y-direction. We have considered an uncertainty of 0.05% for this parameter. This source of uncertainty will be fully correlated between all experiments in the ZPR facility.. In our simulations, the increase of the dimensions is compensated by a decrease in the cell density, so that the total masses of the different components in the system remain constant.

A further limitation is that the input files provided in the benchmarks are homogenized in a small number of regions, which forces us to make some modifications within every one of these regions. In Table 8, it is presented in a summarized form the characteristics considered to be shared and not between the six experiments considered.

Table 7. Main contributors to the uncertainty in k_{eff} in the MCF001-MCF006 benchmark experiments, according to the benchmark documentation (section 2).

	MCF001 -001	MCF002 -001	MCF003 -001	MCF003 -002	MCF004 -001	MCF005 -001	MCF006 -001
TOTAL	**0.08674**	**0.0855**	**0.0687**	**0.0680**	**0.0674**	**0.1116**	**0.0655**
Pu	**?**[8]	**0.0468**	**0.0305**	**0.0285**	**0.0436**	**0.0472**	**0.0446**
U in U_3O_8 (core & reflector)	0.04067	0.04075	---	---	0.0171	---	0.0339
U in DU plates	---	---	0.0421	0.0421	---	0.0854	0.0095
U in Pu-U-Mo fuel	0.02393	0.0131	0.0071	0.0073	0.0137	0.0129	0.0135
Total U	**0.04719**	**0.0428**	**0.0427**	**0.0427**	**0.0219**	**0.0864**	**0.0377**
Matrix tube pitch	**0.04048**	**0.404**	**0.0318**	**0.0318**	**0.0321**	**0.0517**	**0.0276**
Iron & steel (matrix tubes & drawers)	0.0245	0.0235	0.0088	0.0088	0.0083	0.0031	0.0078
Iron & steel (cans, plates & others)	0.0082	0.0032	0.0005	0.0005	0.0047	0.0012	0.0037
Total iron & steel	**0.0258**	**0.0237**	**0.0088**	**0.0088**	**0.0095**	**0.0033**	**0.0086**
U + Pu + pitch + iron	?	0.0788	0.0620	0.0610	0.0592	0.1113	0.0652
% of total	?	92%	90%	90%	88%	>99%	>99%

Table 8. Summary of shared (correlated) and independent (uncorrelated) parameters in the MCF001-MCF006 benchmark experiments.

	MCF001 -001	MCF002 -001	MCF003 -001	MCF003 -002	MCF004 -001	MCF005 -001	MCF006 -001
Pu mass (±0.15 %)	Pu_DOW	Pu_DOW (Pu_21_ic in the inner core)	Pu_SEFOR	Pu_SEFOR (Pu_32_ic in the inner core)	Pu_DOW	Pu_DOW	Pu_DOW
U235 content (±5%)	U_comp_U3O8	U_comp_U3O8	U_comp_UMET	U_comp_UMET	U_comp_U3O8	U_comp_UMET	U_comp_U3O8
U mass (±0.15 %)	U_mass_U3O8	U_mass_U3O8	U_mass_UMET	U_mass_UMET	U_mass_U3O8	U_mass_UMET	U_mass_U3O8
Fe mass (±2%)	Fe_mass	Fe_mass	Fe_mass	Fe_mass	Fe_mass	Fe_mass	Fe_mass
Matrix tuve pitch (±0.05 %)	pitch	pitch	pitch	pitch	pitch	pitch	pitch

[8] This information is omitted in the benchmark documentation (likely typo).

6.2.2 Correlation between integral data

The results of the simulations are presented in Table 9. For comparison, the values proved in DICE are presented in Table 10. The statistical uncertainty, calculated with the formulae obtained in appendix 1, are also listed in the tables. In Figure 5, a couple of examples (with high and low degree of correlation) of the dispersion in the k_{eff} values, from which the correlation coefficient is calculated, are presented.

The first conclusion that is apparent is that the experiments can be classified in two major groups regarding the correlations: one comprising all the experiments with Pu fuel from Dow/Numec (MCF001-001, MCF002-001, MCF004-001, MCF005-001 and MCF006-001) and the other comprising all experiments with Pu fuel from Sefor (MCF003-001 and MCF003-002). These results were to be expected, however, given the assumption that the mass of Pu was fully correlated among all systems sharing the same type of Pu fuel, but uncorrelated between the systems with different types of fuel. Notice that the correlation between these two groups is only due to the steel matrix (Fe mass and matrix pitch), that have been considered to be common for all the seven systems investigated.

Within the first group of systems, MCF005-001 presents somewhat lower correlations with the other systems, which can be attributable to the fact that this is the only system with depleted uranium in the form of U_{MET}, the others having all U_3O_8. At the same time, this can also explain the higher correlations with MCF003-001 and MCF003-002, also featuring U_{MET}. It is also worth remarking the relatively low correlation between MCF004-001 with MCF001-001 and MCF006-001, in spite of these three systems having the same set of correlated parameters.

On the other hand, the effect of replacing the inner core with fuels with higher ^{240}Pu content, and hence considered to be uncorrelated, does not seem to have a relevant impact between the correlations. Indeed, a high level of correlation is observed both between MCF001-001 and MCF002-001 and between MCF003-001 and MCF003-002.

Regarding the comparison with the correlation data available in DICE (as stated above, likely taken from Table XVI of [Palmiotti 2014]), it can be observed that we have found similar values for the correlation between MCF003-001 and MCF003-002 (0.83 vs. 0.85), as well as for the correlations between these systems and MCF004-001 (0.18 vs. 0.21) and between MCF002-001 and MCF003-001/MCF003-002 (0.05 and 0.01 vs 0.07 and 0.06). On the other hand, we have found significantly higher values than the ones listed in DICE for the correlations between MCF001-001 and MCF001-002 (0.87 vs. 0.66) and between MCF001-001/MCF002-001 and MCF004-001 (0.76/0.68 vs. 0.18 and 0.12). In Table XV [Palmiotti 2014], a list of the uncertainty components is given, but no information about the correlations between systems is provided, which does not allow us to investigate the reasons for these discrepancies.

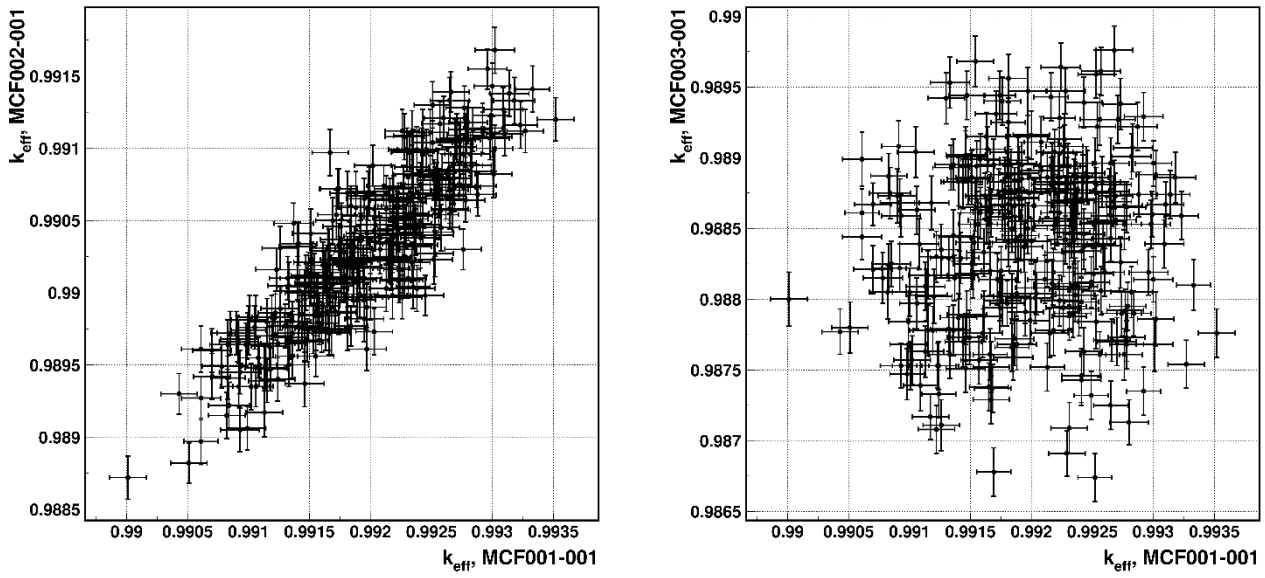
In [Ivanova 2014] and [NEA 2013b], a much lower value for the correlation between MCF001-001 and MCF002-001 is provided (0.134). In [Ivanova 2014], this low value is explained by the assumption of considering as dominant the uncertainty introduced by the homogenization of the calculational model. Once again, this stresses the importance of keeping track of the assumptions made regarding correlated parameters when calculating the experimental corrections between benchmark experiments.

Table 9. Correlations between k_{eff} of the MCF set of benchmark experiments calculated in this work.

		1	2	3	4	5	6	7
1	MCF001-001	1						
2	MCF002-001	0.869 ± 0.010	1					
3	MCF003-001	0.111 ± 0.022	0.050 ± 0.023	1				
4	MCF003-002	0.091 ± 0.023	0.011 ± 0.025	0.829 ± 0.014	1			
5	MCF004-001	0.755 ± 0.015	0.675 ± 0.018	0.180 ± 0.025	0.182 ± 0.026	1		
6	MCF005-001	0.650 ± 0.016	0.575 ± 0.018	0.230 ± 0.022	0.231 ± 0.024	0.772 ± 0.016	1	
7	MCF006-001	0.889 ± 0.009	0.820 ± 0.012	0.124 ± 0.022	0.107 ± 0.024	0.814 ± 0.014	0.725 ± 0.015	1

Table 10. Correlations between k_{eff} of the MCF set of benchmark experiments according to DICE.

		1	2	3	4	5	6	7
1	MCF001-001	1						
2	MCF002-001	0.66	1					
3	MCF003-001	---	0.07	1				
4	MCF003-002	---	0.06	0.85	1			
5	MCF004-001	0.18	0.12	0.21	0.21	1		
6	MCF005-001	---	---	---	---	---	1	
7	MCF006-001	---	---	---	---	---	---	1



(a) MCF002-001 vs. MCF001-001 **(b) MCF003-001 vs. MCF001-001**

Figure 5. Correlations between the k_{eff} of two of the pairs of systems considered in this study. The systems have correlations in all parameters considered in Table 8, except the fuel in the inner core. The systems on the right only have correlations in the iron mass and the matrix tube pitch.

6.3 ASPIS Iron-88 shielding benchmark

The ASPIS Iron-88 benchmark experiment [Wright 1993, Avery 1995, Milocco, 2015] was performed in 1988 in the ASPIS shielding facility installed next to the Nestor reactor at Winfrith to study the neutron transport for penetrations up to 67 cm in steel. The experimental array irradiated in ASPIS comprises a fission plate made of 93% enriched U-Al alloy driven by thermal neutrons from the NESTOR reactor and installed in front of the shield made from 13 mild steel plates each 5.1 cm thick, and a deep backing shield manufactured from mild and stainless steel. Absolute source strength and spatial distribution were determined by fission product counting and ^{55}Mn(n,γ) measurements over the X-Y front surface. However, the uncertainties of these measurements were not reported. Au, Rh, In, S and Al activation foils were placed in ~7.4-mm air gaps between each slab component along the fission plate axis at several shield thicknesses up to ~67 cm.

Detailed information on the systematic and statistical uncertainties of the measurements was reported by the experimentalists and is shown in Table 11 [Wright 1993, Avery 1995]. Systematic and statistical uncertainties are well separated and characterised in the reports and allowed to construct the correlation matrix of the measured reaction rates partly presented in Table 12 (only a selected subset is shown for demonstration here). On the other hand, no information is available on the background correction uncertainty and thermal flux leaking from the Nestor reactor, which may be non-negligible for the gold foil measurements, uncertainties in in the foil positioning and arrangement, and in the geometry descriptions and modelling, material compositions and dimensions.

Detector	**Diameter** (mm)	**Thickness** (mm)	**Mass** (g)	**Uncertainty (%)** Systematic calibration	Counting statistics	Power	Total
^{197}Au(n,γ)	12.7	0.05	0.12-0.13	0.9	1	4	4.2
^{103}Rh(n,n')	12.7	0.015	0.20	3.0	1 - 1.5	4	5.1 - 5.2
^{115}In(n,n')	38	1.63	12.79	1.9	1 - 1.7	4	4.5 - 4.7
^{32}S(n,p) pressed pellet	38.1	2.41	5	5.0	1 - 1.5	4	6.5
^{32}S(n,p) cast pellet	51	5.6	22	5.0	1 - 5.7 (up to 20)	4	6.5 - 8.6 (up to 21)
^{27}Al(n,α)	50	3.1	16.72	2.2	1 - 1.3	4	4.7

Table 11. Components of the measurement uncertainties (1σ) for the reactions measured in ASPIS Iron-88 benchmark.

In the evaluation of the covariance matrix of these measurements presented in Table 12 [Kodeli 2018], the following assumptions were made:

- Power normalisation uncertainty is assumed to be correlated over all reaction rates. The justification for this assumption is though partly questionable since there is no evidence that all the foils were irradiated in a single or in several experimental campaigns;
- Calibration uncertainty is a component correlated over all detector positions, but uncorrelated with other reaction rates. In practice, there may be some correlation between the calibration uncertainties of the different detectors;
- Statistical uncertainty component is (of course) the uncorrelated component of the uncertainty.
- Other sources of uncertainty than in the calibration, power and counting statistics were neglected.

In addition, the covariance matrix relative to the ratios of the measured reaction rates with respect to the 1st measurement position was prepared for the purpose of the WPEC WG39 activities [Salvatores 2014, WPEC39]. The matrix is presented in Table 13 for the same subset of detectors/foil positions as above. It can be observed that, with the four above assumptions, the use of reaction rate ratios significantly reduces the uncertainties and un-correlates the covariance matrix. A drawback is that dividing the reaction rate of each detector by the 1st reaction rate position of the same reaction means that the effective normalisation factor becomes different for each detector and reaction. This may lead to the loss of information on the spectra distribution. In the case of ASPIS Iron-88 the effect is particularly strong for the Al(n,α) reaction with the observed C/E values around 1.35. Adjustment based on the Table 13 type of covariances may therefore not be able to correct for the possible/probable inconsistencies for this high energy reaction. The four assumptions may be also not entirely realistic for the adjustment applications.

			Au			Rh		In		S			Al
	Pos.		A7	A11	A14	A7	A14	A7	A11	A7	A12	A14	A7
		1σ(%)	4.2	4.2	4.2	5.1	5.1	4.5	4.7	6.5	6.5	8.6	4.7
Au	A7	4.2	1.00	0.944	0.944	0.744	0.744	0.835	0.799	0.585	0.585	0.442	0.799
	A11	4.2		1.00	0.944	0.744	0.744	0.835	0.799	0.585	0.585	0.442	0.799
	A14	4.2			1.000	0.744	0.744	0.835	0.799	0.585	0.585	0.442	0.799
Rh	A7	5.1				1.000	0.962	0.691	0.662	0.484	0.484	0.366	0.661
	A14	5.1					1.000	0.691	0.662	0.484	0.484	0.366	0.661
In	A7	4.5						1.000	0.911	0.544	0.544	0.411	0.743
	A11	4.7							1.000	0.520	0.520	0.393	0.711
S	A7	6.5								1.000	0.976	0.738	0.520
	A12	6.5									1.000	0.738	0.520
	A14	8.6										1.000	0.393
Al	A7	4.7											1.000

Table 12. ASPIS Iron-88 covariance matrix for the measured reactions rates, including the power normalization, detector calibration and statistical uncertainties. The covariance matrix is constructed so as to contain the total standard deviation (including systematic plus stochastic uncertainties) in the diagonal, and the systematic part of the uncertainty in the off-diagonal positions. The power normalization uncertainty was assumed to be completely correlated among the detectors. Note that only the upper part of the (symmetric) matrix is given.

			Au			Rh		In		S		
	Pos.		A7 /A2	A11 /A2	A14 /A2	A7 /A2	A14 /A2	A7 /A2	A11 /A2	A7 /A2	A12 /A2	A14 /A2
		1σ(%)	5.9	5.9	5.9	7.2	7.2	6.4	6.5	1.4	1.4	5.8
Au	A7/A2	1.4	1.00	0.50	0.50	0	0	0	0	0	0	0
	A11/A2	1.4		1.00	0.50	0	0	0	0	0	0	0
	A14/A2	1.4			1.00	0	0	0	0	0	0	0
Rh	A7/A2	1.4				1.00	0.50	0	0	0	0	0
	A14/A2	1.4					1.00	0	0	0	0	0
In	A7/A2	1.4						1.00	0.49	0	0	0
	A11/A2	2.0							1.00	0	0	0
S	A7/A2	1.4								1.00	0.50	0.12
	A12/A2	1.4									1.00	0.12
	A14/A2	5.8										1.00

Table 13. ASPIS Iron-88 covariance matrix for the ratios of the measured reactions rates. Note that the correlated components of the uncertainty, such as the power normalization, detector calibration uncertainties, were supposed to cancel out. Note that only the upper part of the (symmetric) matrix is given.

6.4 VENUS-3 shielding benchmark (power distribution)

The VENUS-3 benchmark [Hondt 1990, Leenders 1988] was designed essentially to test refuelling patterns reducing the radiation exposure to the reactor pressure vessel and the ability of the fluence-rate synthesis procedures used at the time based on 2D/1D calculations to predict adequately 3D geometry effects.

The VENUS-3 power distribution was fully measured only at two (out of 14) axial levels [Abderrahim 1999], namely at the axial level corresponding to the mid-plane of the lower-PLSA (Partial Length Shielded Assembly) part of the core loading, and at the one corresponding to the mid-plane of the upper part of the core loading. In addition to the two XY radial distributions, the full axial power distribution was measured at 374 fuel pin locations, out of the total of 639 fuel pins comprised in the 1/4 of the reactor core.

In order to establish a complete 3D map of the power distribution in the VENUS core, an extrapolation procedure, based on the RECOG-ORNL code [Begovich 1977] was used back in 1997 when the VENUS-3 benchmark was evaluated for the purpose of an international inter-comparison exercise and the SINBAD database. RECOG-ORNL is a general-purpose pattern recognition code. Various methods for data analysis, pre-processing and display of data, unsupervised and supervised learning are available in the code. Hence, the data provided in this benchmark contain the measured values, where available, and RECOG predicted values elsewhere. Relative nuclear power distribution, normalised to the core averaged power of one fission per second per active pin, are given for each fuel pin position and for 14 axial levels (see examples in Figure 6).

This procedure permitted at the same time to detect some suspicious or faulty values, transcription errors, as well as to give an idea of the accuracy of the neutron flux. This accuracy was evaluated in the following way: where the measured values were available the 'uncertainties' were determined as the % difference between the neutron source values calculated by RECOG code and the measured values. Elsewhere the uncertainty of the extrapolation procedure in RECOG was estimated to ±5% [NEA SINBAD].

To verify the results obtained at the time and to test the possible use and performances of the new ML tools, the interpolation procedure has been recently re-evaluated using modern ML methods [Berger 2024]. Different algorithms have been considered, such as Pattern recognition, Linear regression, Lasso regression, Ridge regression, Kriging, K-Nearest Neighbours (KNN), Support Vector Machine, Random Forest and Neural Networks (NN, or Artificial Neural Network, ANN). Predictions of different algorithms are compared in terms of performance parameters to obtain conclusions on the method performances. The following performance criteria have been used:

- Mean absolute error in %

$$MAE = \frac{100\%}{N} \sum_{i=1}^{N} \left| \frac{y_{exp,i} - \vec{y}_i}{y_{exp,i}} \right| \qquad (6.1)$$

- Mean square error (MSE)

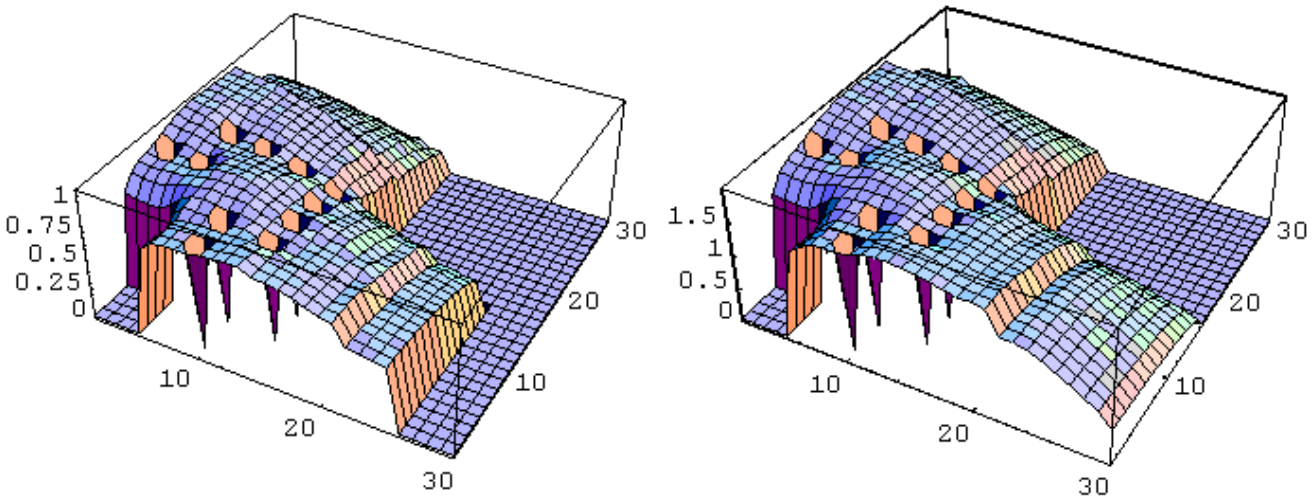
$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_{exp,i} - \vec{y}_i\ )^2 \qquad (6.2)$$

- Prediction coefficient

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_{exp,i} - \vec{y}_i\ )^2}{\sum_{i=1}^{N} (\overline{y_{exp}} - \vec{y}_i\ )^2} \qquad (6.3)$$

The accuracy of the power map in the reactor core and the uncertainties of the extrapolation procedure are compared for the different ML methods used. Furthermore, as a key objective, the correlations between the measurements, and, in particular, the correlations among the corresponding uncertainties are studied using the predictions of different fitting algorithms. Predictions of different algorithms are compared and allow to evaluate the method performances in terms of nominal values, average absolute and quadratic error, prediction coefficients, residues, and covariance matrices. First preliminary results are presented in Table 14 and Table 15,

and Figure 7 and Figure 8. The results indicate good general performances of some of the methods (except linear regression, Polynomial and Neural Networks) and a progress compared to the Pattern recognition algorithm used in the past. As shown in Table 15 the inter-/extrapolated values calculated using the Kriging, KNN, Gradient Boosting and Random Forest methods agreed well (on the average within less than 1%) with the predictions obtained using the Pattern Recognition technique. The work on the elaboration and testing of the covariance matrix is still ongoing and the results are encouraging.



(a) Axial level 1 **(a) Axial level 2**

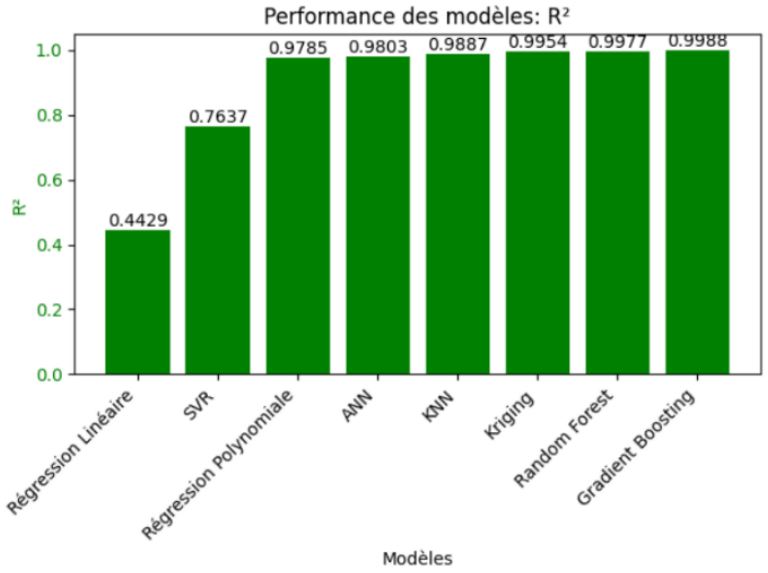
Figure 6. VENUS-3 radial relative power distributions at two axial levels 1 and 9.

Modèle(%)	Moyenne(%)	Maximum(%)	Minimum(%)
RECOG-ORNL	3.15	23.5	0.1
Kriging	0.57	36.6	3.3e-05
KNN	0.61	27.1	0.0
Gradient Boosting	0.73	24.4	7.48e-06
Random Forest	0.76	16.2	1.9e-14
ANN	4.33	56.1	1.4e-03
Polynomial	4.37	110.6	6.1e-4
Linéaire	23.3	67.1	3.8e-3

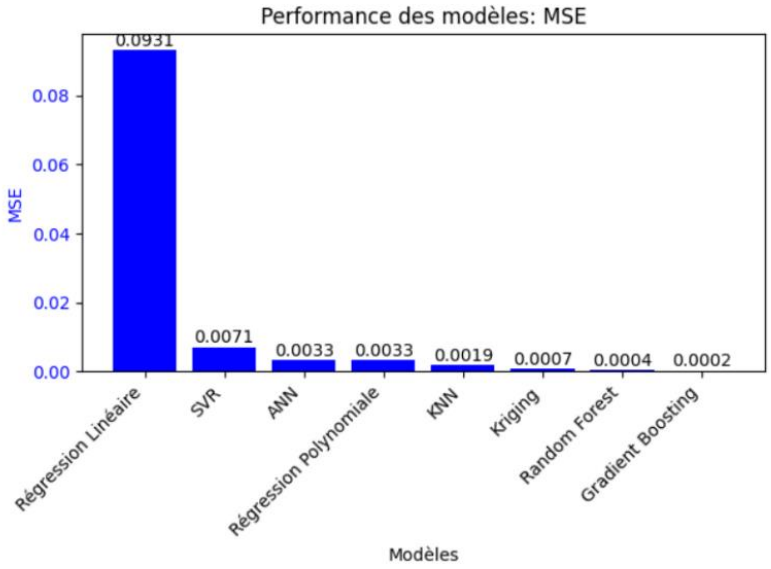
Table 14. Summary of the performances of different ML models in terms of mean, maximum and minimum absolute errors in %.

Modèle(%)	Moyenne(%)	Maximum(%)	Minimum(%)
Kriging	0.48	12.9	4.02e-05
KNN	0.66	33.9	0.0
Gradient Boosting	0.69	14.1	7.5e-06
Random Forest	0.78	16.8	8e-04
ANN	4.41	27.3	1.4e-03
Linéaire	26.6	185.5	0.01
Polynomial	37.5	37.6	5.8e-03

Table 15. Mean, maximal and minimal differences in the predicted power distribution of different ML models with respect to the Pattern Recognition method.

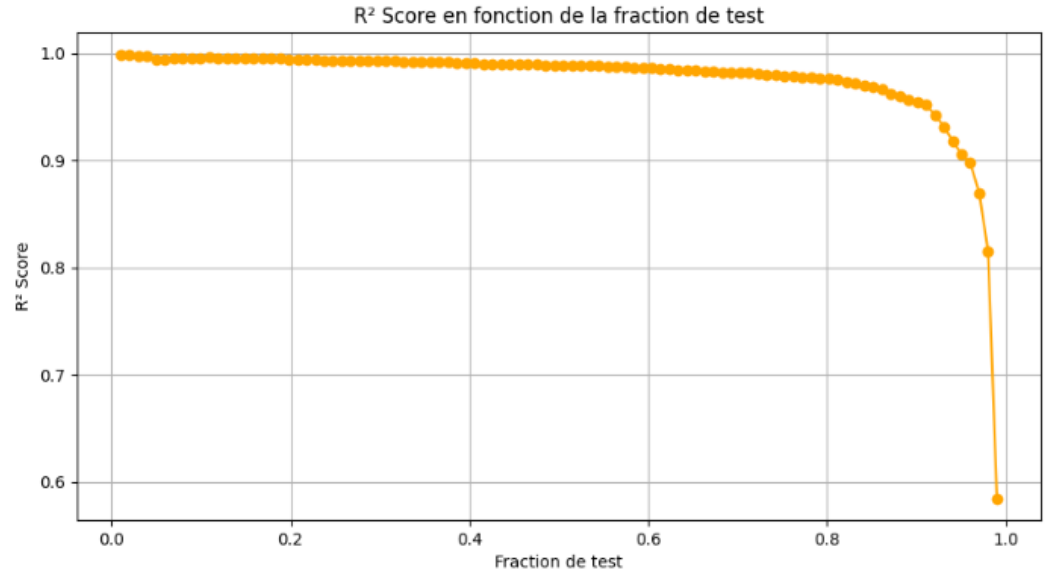


(a) R^2

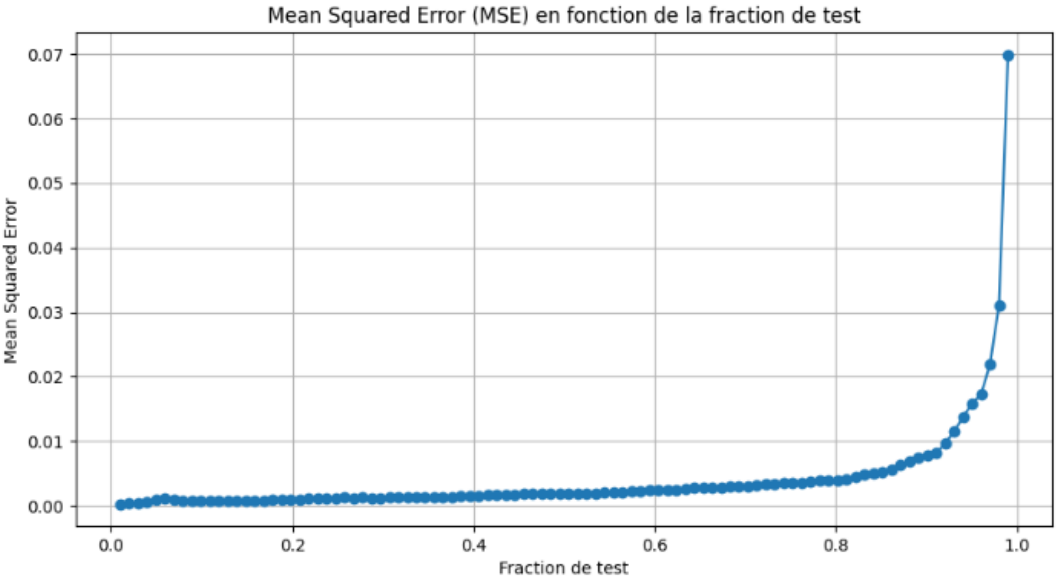


(b) Mean square error (MSE)

Figure 7. Prediction parameter R^2 and Mean square error (MSE) for different ML methods.



(a) R^2



(b) Mean square error (MSE)

Figure 8. Evolution of the Prediction parameter R^2 and Mean square error (MSE) with the fraction of training sets used (Gradient Boosting method).

7 Summary & conclusions

Correlations between the uncertainties in experimental parameters of criticality benchmark experiments (and other reactor physical parameters) are known to exist and to play an important role in some important applications, such as criticality safety assessment and nuclear data adjustment. Although methodologies to determine these correlations have been developed, both deterministic and Monte Carlo, no systematic and comprehensive study of these correlations has been performed up to now. The most important source presently available are the data included in OECD/NEA DICE database, and it contains information about experimental correlations only for 93 cases out of the more than 5000 included in the ICSBEP database. In this work, we present the results of some additional calculations: between two configurations of the EOLE reactor (calculated with the deterministic methodology) and between seven configurations of the ZPR reactor (MIX-COMP-FAST) included in the ICSBEP database (calculated with the Monte Carlo methodology). Furthermore, an example of the evaluation of the correlation matrix for a set of activation-foil measurements carried out during a shielding experiment included in the SINBAD database (ASPIS Iron-88) is also provided.

The users have to be aware that the evaluation of the experimental correlations is to a large extent based on expert judgment (both with the deterministic and Monte Carlo methodologies), and information about the hypotheses made in this respect is not always accessible. Furthermore, it is often difficult to extract reliable information on the systematic and stochastic uncertainties for the older experiments. In particular, the DICE database includes no information about the sources of the experimental correlation values listed in it, which makes it very difficult to reproduce these values. Extreme caution is therefore required in the use of this information in e.g. data adjustment analysis. Using the ratios of strongly correlated data (such as reaction rates, k_{eff}, etc.) largely reduces or even eliminates the systematic uncertainties and can lead to considerable reduction in total uncertainties.

Another important aspect to take into account, when applying the Monte Carlo methodology to determine the experimental correlations, is that statistical uncertainties have to be calculated and reported. A formula to propagate the uncertainties in two sets of statistical variables to their correlation coefficient has been obtained and is presented in Appendix 1.

Finally, an example (VENUS-3 shielding benchmark) of the potential benefits of the use of machine learning techniques in the interpretation of experimental uncertainties and correlations is also provided in this work. Physical interpretation of the results obtained is a major challenge when applying these techniques.

Acknowledgements

This work has been funded by the Euratom research and innovation programme under the SANDA project (Grant Agreement No. 847552). The ICSBEP database has been obtained through the Computer Programs Services of OECD/NEA Nuclear Data Bank and the MCNP 6.2 code through the Radiation Safety Information Computational Center (RSICC) of Oak Ridge National Laboratory (ORNL, USA). We are also grateful to Juan Blázquez Martínez (CIEMAT) for the careful revision of the manuscript.

References

[Abderrahim 1999] H. Ait Abderrahim, *VENUS-3 3-D Power Distribution*. Private communication.

[Avery 1995] A. Avery and I. Kodeli, *Winfrith Iron 88 Benchmark (ASPIS)*. SINBAD package (NEA-1517/35).

[Begovich 1977] C. L. Begovich and N. M. Larson, *A User's Manual for the Pattern Recognition Code RECOG-ORNL*. ORNL/CSD/TM-21 (1977).

[Berger 2024] Adèle Berger, Ivan Kodeli and Pierre-Jacques Dossantos-Uzarralde, *Application of Machine Learning techniques to the exploitation and interpolation of nuclear experimental data*. Nuclear Energy for New Europe Conference (NENE-2024), Portorož (Slovenia) 9-12 September 2024.

[Buss 2010] O. Buss *et al.*, *Hierarchical Monte-Carlo approach to bias estimation for criticality safety calculations*. Proceedings of PHYSOR2010, Pittsburgh (PA, USA) 9-14 May 2010.

[Dean 2008] V.F. Dean (Ed), *ICSBEP Guide to the Expression of Uncertainties*. https://www.oecd-nea.org/jcms/pl_24501/international-criticality-safety-benchmark-evaluation-project-icsbep-guide-to-the-expression-of-uncertainties?details=true

[Dos Santos 2013a] N. dos Santos *et al.*, *Impact of Mock-up Experimental Correlations and Uncertainties in the Transposition Process*, Proceedings of ANIMMA 2013, Marseille (France) 23-27 June 2013.

[Dos Santos 2014] N. dos Santos, *Optimisation de l'approche de représentativité et de transposition pour la conception neutronique de programmes expérimentaux dans les maquettes critiques*, PhD thesis, university of Grenoble (France), 16 april 2014. https://theses.hal.science/tel-00979526.

[García-Herranz 2024] N. García-Herranz, O. Cabellos, A. Jiménez-Carrascosa, D. Bernard, I. Kodeli, G. Zerovnik and S. van der Marck, *Report on reactor and shielding C/E validation and nuclear data trends*, SANDA Project Deliverable D5.7 (18 June 2024).

[GRS 2016a] M. Stuke (Ed.), *Proceedings of the Workshop on Integral Experiment Covariance Data for Critical Safety Validation.* Technical report GRS-414, Gesellschaft für Anlagen und Reaktorsicherheit (GRS) GmbH, April 2016. ISBN 978-3-944161-96-9.

[GRS 2016b] M. Stuke *et al.*, *Generation of Integral Experiment Covariance Data and their Impact on Criticality Safety Validation*. Technical report GRS-440, Gesellschaft für Anlagen und Reaktorsicherheit (GRS) GmbH, November 2016. ISBN: 978-3-946607-22-9.

[Hondt 1990] P. D'Hondt *et al.*, *Contribution of the VENUS-Engineering Mock-up Experiment to the LWR-PV-Surveillance*. Proc. 7th ASTM-EURATOM Symposium on Reactor Dosimetry, Strasbourg (France) 27-31 August 1990.

[Ivanova 2003a] T. T. Ivanova *et al.*, *Influence of the correlations of experimental uncertainties on criticality prediction*. Nuclear Science and Engineering 145 (2003) 97-104. DOI: 10.13182/NSE03-27.

[Ivanova 2003b] T. T. Ivanova *et al.*, *Attempt of the Joint Analysis of the Entire Set of the HEU-SOL Type Experiments from the "International Handbook of Evaluated Criticality Safety Benchmark Experiments"*. Proceedings of the 7th International Conference on Nuclear Criticality Safety (ICNC 2003). Tokai (Ibaraki, Japan) 20-24 October 2003. DOI: 10.11484/JAERI-Conf-2003-019-Part2.

[Ivanova 2009] T. T. Ivanova *et al.*, *Towards validation of criticality calculations for systems with MOX powders*. Annals of Nuclear Energy 36 (2009) 305-309. DOI: 10.1016/j.anucene.2008.12.018.

[Ivanova 2014] T. T. Ivanova *et al.*, *Establishment of Correlations for Some Critical and Reactor Physics Experiments*. Nuclear Science and Engineering 178 (2014) 1–15.

[Jeong 2017] S. Jeong *et al.*, *Creation of a Database of Uncertainties for ICSBEP Handbook and Tool for Covariance Generation*. Proceedings of M&C 2017, Jeju (Korea) 16-20 April 2017.

[Kilger 2016] R. Kilger *et al.*, *Analysis of experimental series of plutonium nitrate in aqueous solution and their correlation coefficients*, Annals of Nuclear Energy 96 (2016) 354–362. DOI: 10.1016/j.anucene.2016.06.017.

[Kodeli 2014] I. Kodeli, A. Milocco, P. Ortego and E. Sartori, *20 Years of SINBAD (Shielding Integral Benchmark Archive and Database)*, Progress in Nuclear Science and Technology 4 (2014) 308-311.

[Kodeli 2018] I. Kodeli and L. Plevnik*, Nuclear data adjustment exercise combining information from shielding, critical and kinetics benchmark experiments ASPIS-Iron 88, Popsy and SNEAK-7A/7B*, Progress in Nuclear Energy 106 (2018) 215 – 230.

[LANL 2017] C. J. Werner (Ed.), *MCNP User's Manual Code Version 6.2*. Technical Report LA-UR-17-29981, Los Alamos National Laboratory (27 October 2017).

[Leenders 1988] L. Leenders, *LWR-PVS Benchmark Experiment VENUS-3*. FCP/VEN/01 report (September 1988).

[Marshall 2015] W. J. Marshall and B. T. Rearden, *Determination of Critical Experiment Correlations Using the Sampler Sequence Within SCALE 6.2*. Proceedings of 10th International conference on Nuclear Criticality Safety (ICNC 2015), Charlotte (North Carolina, USA) 13-17 September 2015.

[Marshall 2017a] W. J. Marshall *et al.*, *Determination of critical experiment correlations for experiments involving arrays of low-enriched fuel rods*. ANS Nuclear Criticality Safety Division Topical Meetings (NCSD2017), Carlsbad (New Mexico, USA) 10-15 September 2017. ISBN: 978-0-89448-740-8.

[Marshall 2017b] W. J. Marshall *et al.*, *Determination of critical experiment correlations for experiments involving highly enriched uranium solutions*. ANS Nuclear Criticality Safety Division Topical Meetings (NCSD2017), Carlsbad (New Mexico, USA) 10-15 September 2017. ISBN: 978-0-89448-740-8.

[Marshall 2017c] W. J. Marshall, Determination of critical experiment correlations via the Monte Carlo sampling technique, PhD Thesis, University of Tennessee (USA), December 2017. https://trace.tennessee.edu/utk_graddiss/4750.

[Marshall 2019] W. J. Marshall *et al.*, *Correlation of HST-001 due to uncertain technical parameters – Comparison of results from DICE, SAMPLER and SUnCISTT*. Proceedings of 11th International conference on Nuclear Criticality Safety (ICNC 2019), Paris (France) 15-20 September 2019.

[Milocco 2015] A. Milocco, *Quality Assessment of SINBAD Evaluated Experiments ASPIS Iron (NEA-1517/34), ASPIS Iron-88 (NEA-1517/35), ASPIS Graphite (NEA-1517/36), ASPIS Water (NEA-1517/37), ASPIS N/G Water/Steel (NEA-1517/49), ASPIS PCA Replica (NEA-1517/75)*. (December 2015).

[NEA 2010] C. de Saint-Jean *et al.*, *Assessmet of Existing Nuclear Data Adjustment Methodologies*. Technical Report NEA/NSC/WPEC/DOC(2010)429, OECD/NEA, 2010.

[NEA 2013a] T. T. Ivanova *et al.*, *Overview of Approaches Used to Determine Calculational Bias in Criticality Safety Assessment*. Technical Report NEA/NSC/WPEC/DOC(2013)7, OECD/NEA, October 2013.

[NEA 2013b] M. Salvatores and G. Palmiotti [Coords.] Methods and Issues for the Combined Use of Integral Experiments and Covariance Data. Technical Report NEA/NSC/WPEC/DOC(2013)445, OECD/NEA, December 2013.

[NEA 2023] *Role of Integral Experiment Covariance Data for Criticality Safety Validation. EG UACSA Benchmark Phase IV*. Technical Report NEA/NSC/R(2021)1, OECD/NEA, January 2023.

[NEA DICE] *Database for ICSBEP (DICE)* (database). https://www.oecd-nea.org/jcms/pl_20293/database-for-icsbep-dice.

[NEA ICSBEP] *International Handbook of Evaluated Criticality Safety Benchmark Experiments* (database). https://www.oecd-nea.org/jcms/pl_20291/icsbep-handbook (Accessed 26 April 2024).

[NEA IRPhE] *International Handbook of Evaluated Reactor Physics Benchmark Experiments* (database). OECD/NEA. https://www.oecd-nea.org/jcms/pl_20279/international-handbook-of-evaluated-reactor-physics-benchmark-experiments-irphe (Accessed 26 April 2024).

[NEA SINBAD] SINBAD - Shielding Integral Benchmark Archive and Database (database). OECD/NEA. https://www.oecd-nea.org/jcms/pl_32139/sinbad-shielding-integral-benchmark-archive-and-database

[NEA WPEC39] Working Party on Evaluation Cooperation (WPEC), *Methods and approaches to provide feed-back from nuclear and covariance data adjustment for improvement of nuclear data files*, Interna-ional Evaluation Co-operation, NEA/WPEC-39, OECD/NEA, ongoing (https://www.oecd-nea.org/science/wpec/sg39/).

[NJOY 2021] NJOY21—NJOY for the 21st century. URL https://www.njoy21.io/NJOY21/. (Accessed 27 April 2023).

[Nouri 2003] A. Nouri *et al.*, *DICE: Database for the International Criticality Safety Benchmark Evaluation Program Handbook*. Nuclear Science and Engineering 145 (2003) 11-19. DOI: 10.13182/NSE03-15.

[Palmiotti 2014] G. Palmiotti *et al.*, *Combined Use of Integral Experiments and Covariance Data*. Nuclear Data Sheets 118 (2014) 596–636. DOI: 10.1016/j.nds.2014.04.145.

[Peters 2016] E. Peters *et al.*, *Modeling of critical experiments and its impact on integral covariance matrices and correlation coefficients*. Annals of Nuclear Energy 92 (2016) 355–362. DOI: 10.1016/j.anucene.2016.02.011.

[Plompen 2020] A. J. M. Plompen *et al.*, 2020. *The joint evaluated fission and fusion nuclear data library, JEFF-3.3*. Eur. Phys. J. A 56 (2020). 1–108. DOI: 10.1140/epja/s10050-020-00141-9.

[Salvatores 2014] M. Salvatores *et al.*, *Methods and Issues for the Combined Use of Integral Experiments and Covariance Data: Results of a NEA International Collaborative Study*. Nuclear Data Sheets, 118 (2014) 38-71. DOI: 10.1016/j.nds.2014.04.005.

[Santamarina 1982] A. Santamarina and C. Cerdan, *Expérience CAMELEON : Description du cœur pour les calculs de neutronique, géométrie et bilan matière*, Technical note CEA/SEN/LPRE/82-203 (1982).

[Schmelling 1995] M. Schmelling, *Averaging Correlated Data*. Physica Scripta 51 (1995) 676-679. DOI: 10.1088/0031-8949/51/6/002.

[Sommer 2021] F. Sommer and M. Stuke, *An efficient and fast method to calculate integral experimental correlation coefficients – S^2Cor*. Annals of Nuclear Energy 157 (2021) 108209. DOI: 10.1016/j.anucene.2021.108209.

[Stuke 2019] M. Stuke *et al.*, *UACSA Phase IV: role of integral experiment covariance data for criticality safety validation. Summary of selected results*. Proceedings of 11th International conference on Nuclear Criticality Safety (ICNC 2019), Paris (France) 15-20 September 2019.

[Tabashi 2018] M. Tanabashi *et al.* (Particle Data Group), *Review of Particle Physics*. Physical Review D 98 (2018) 030001. DOI: 10.1103/PhysRevD.98.030001.

Appendix 1. Propagation of statistical errors in the MC methodology

The Pearson's correlation coefficient of two sets of data (x_1, …, x_n) and (y_1, …, y_n) is defined as:

$$\rho_{xy} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}} \quad \text{(A.1)}$$

If we assume that both data sets x_i and y_i are affected by uncertainties, and that these uncertainties are independent (i.e. non-correlated), then it is possible to apply the the classical first-order error propagation formula to obtain the error in ρ_{xy}:

$$\delta\rho_{xy} = \sqrt{\left(\frac{\partial\rho_{xy}}{\partial x_i}\right)^2 \delta^2 x_i + \left(\frac{\partial\rho_{xy}}{\partial y_i}\right)^2 \delta^2 y_i} \quad \text{(A.2)}$$

To calculate the derivatives of ρ_{xy}, let us write it in terms of the three following functions:

$$f_{xy}(x_1, \dots, x_n, y_1, \dots, y_n) = n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i \quad \text{(A.3)}$$

$$g_x(x_1, \dots, x_n) = \sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad \text{(A.4)}$$

$$g_y(y_1, \dots, y_n) = \sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2} \quad \text{(A.5)}$$

Hence:

$$\rho_{xy} = \frac{f_{xy}(x_1,\dots,x_n,y_1,\dots,y_n)}{g_x(x_1,\dots,x_n)g_y(y_1,\dots,y_n)} \quad \text{(A.6)}$$

$$\frac{\partial\rho_{xy}}{\partial x_i} = \frac{\frac{\partial f_{xy}}{\partial x_i}g_x - f_{xy}\frac{\partial g_x}{\partial x_i}}{g_x^2 g_y} \quad \text{(A.7)}$$

$$\frac{\partial\rho_{xy}}{\partial y_i} = \frac{\frac{\partial f_{xy}}{\partial y_i}g_y - f_{xy}\frac{\partial g_y}{\partial y_i}}{g_y^2 g_x} \quad \text{(A.8)}$$

The derivatives of every one of these functions are:

$$\frac{\partial f_{xy}}{\partial x_j} = \frac{\partial}{\partial x_j}\left(n\sum_{i=1}^{n} x_i y_i - \sum_{j=1}^{n} x_i \sum_{i=1}^{n} y_i\right) = ny_j - \sum_{i=1}^{n} y_i \quad \text{(A.9)}$$

$$\frac{\partial g_x}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\right) = \frac{2nx_j - 2\sum_{i=1}^{n} x_i}{2\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}} = \frac{nx_j - \sum_{i=1}^{n} x_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}} \quad \text{(A.10)}$$