

# Improving neutronics simulations and uncertainties via a selection of nuclear data<sup>\*</sup>

D. Rochman<sup>1,a</sup>, A.J. Koning<sup>2</sup>, and S.C. van der Marck<sup>2</sup>

<sup>1</sup> Laboratory for Reactor Physics Systems Behaviour, Paul Scherrer Institut, Villigen, Switzerland

<sup>2</sup> Nuclear Research and Consultancy Group NRG, Petten, The Netherlands

Received: 11 December 2014 / Revised: 6 March 2015

Published online: 24 December 2015 – © Società Italiana di Fisica / Springer-Verlag 2015

Communicated by F. Gulminelli

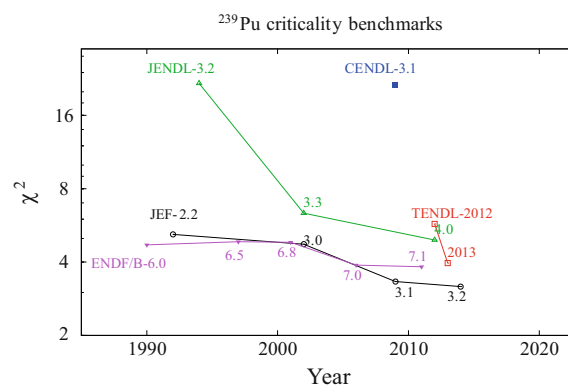
**Abstract.** This work presents a novel approach to improve neutronics simulations, as in the case of criticality calculations, by simply combining the results of a limited set of random evaluations. Another outcome of this work is to lower uncertainties due to nuclear data by integrating the information from criticality benchmarks into the neutronics simulations scheme. Examples are presented for the <sup>239</sup>Pu nuclear data and calculations of criticality benchmarks and a MOX fuel pincell.

## 1 Introduction

The goal of a nuclear data evaluator is to provide the nuclear community with evaluated quantities (best estimates) which are in agreement with measurements and theoretical expectations. The ultimate achievement is that these evaluated quantities are in perfect agreement with all measurements and theoretical calculations. It is of course virtually impossible in practice since measurements can be inconsistent and theories are very often approximate. Therefore a more feasible achievement is to provide evaluated quantities presenting the best possible agreement with all measurements and theoretical calculations. This is possible in practice and nuclear data evaluators devote a lifetime of efforts to produce such data. To realize this goal, the evaluators can use a series of differential and integral measurements, the difference being that the differential measurements are aiming to provide information on a single type of nuclear data, such as the fission cross section of <sup>239</sup>Pu, whereas integral measurements can indirectly give indications on a set of specific nuclear data. This last type of measurement can also be used to quantify the quality of a library (collection of nuclear data for different isotopes) in connection with a simulation tool. In the following, we will restrict ourselves to criticality measurements of  $k_{\text{eff}}$ , which represents one of the most extensive collections of integral data [1], and we will use a very simple (but effective) quantity to compare the performance of libraries: the  $\chi^2$  value per point as defined by

<sup>\*</sup> Contribution to the Topical Issue “Perspectives on Nuclear Data for the Next Decade” edited by Nicolas Alamanos, Eric Bauge, Stéphane Hilaire.

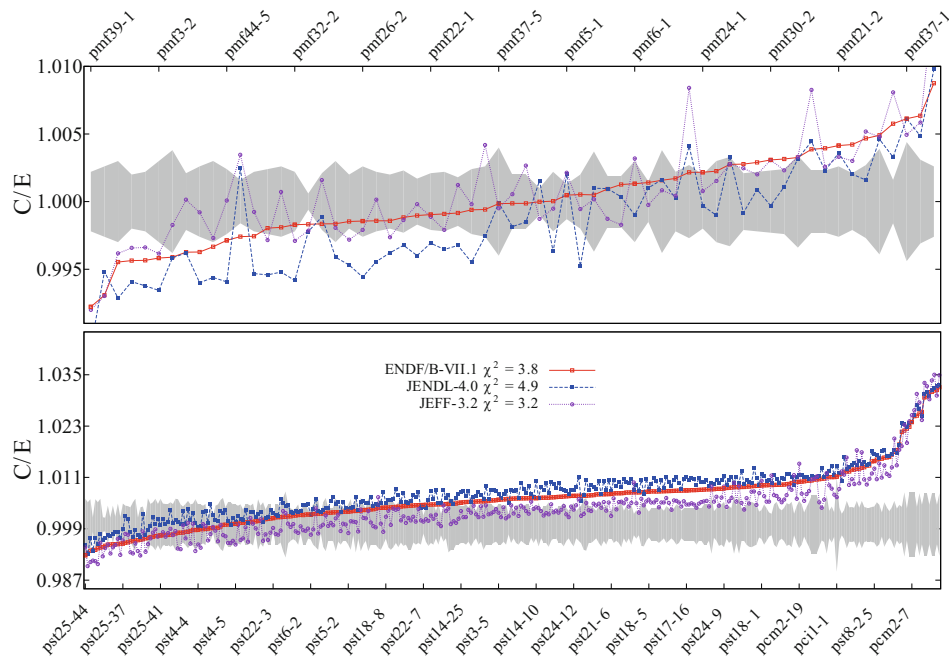
<sup>a</sup> e-mail: dimitri-alexandre.rochman@psi.ch



**Fig. 1.**  $\chi^2$  values as defined in eq. (1) for a series of nuclear data libraries: ENDF/B for the USA [3], JENDL for Japan [4], CENDL for China [5], JEF(F) for the OECD/NEA [6] and TENDL for the TALYS-based library [7]. 400 <sup>239</sup>Pu criticality benchmarks are taken into account (among which 60 defined as *fast* benchmarks). The numbers indicated close to each point are the library release version.

$$\chi^2 = \frac{1}{n} \sum_{i=1}^n \left( \frac{C_i - E_i}{\Delta E_i} \right)^2, \quad (1)$$

where  $C_i$  is the calculated  $k_{\text{eff}}$  (using the MCNP simulation tool [2]),  $E_i$  is the measured  $k_{\text{eff}}$  as indicated in ref. [1],  $\Delta E_i$  is the experimental uncertainty and  $n$  is the number of considered benchmarks. If a library presents a better performance than another one, then its  $\chi^2$  value is smaller. To illustrate the values  $\chi^2$  can take, an example for 400 <sup>239</sup>Pu criticality benchmarks is presented in fig. 1 (60 fast benchmarks and 340 mixed and thermal),



**Fig. 2.** Status of the criticality benchmarks for  $^{239}\text{Pu}$  for 400 criticality benchmarks and for 3 main libraries: JEFF-3.2, JENDL-4.0 and ENDF/B-VII.1. Top: fast benchmarks. Bottom: all others. The benchmarks in the  $x$ -axis are ordered from the smallest  $C/E$  to the highest one for ENDF/B-VII.1.

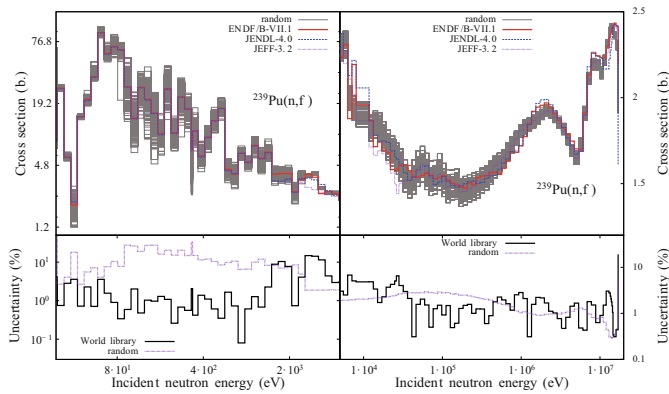
obtained using MCNP and different releases of 5 nuclear data libraries. In fig. 1, the  $\chi^2$  values for each different release per library are showing a global improvement over time, partially confirming the above statement that evaluators provide the best possible agreement with integral values (at a given time). For indication, the JEFF-3.2 library has a  $\chi^2$  value of 3.2, whereas the JENDL-4.0 library has a value of 4.9; assuming an average of 400 pcm for all  $\Delta E_i$ , the average deviation for JEFF-3.2 is 1300 pcm and 2000 pcm for JENDL-4.0. In the case of fast benchmarks only, these values are reduced to 650 and 800 pcm, respectively. It is interesting to remark that even if the different releases of a library are showing a global decrease in  $\chi^2$ , the latest releases present a modest improvement. This characteristics should not be confounded with the general quality of a library which contains much more information than described by a  $\chi^2$  value for  $^{239}\text{Pu}$  criticality benchmarks.

From fig. 1, one can look more closely at the performance of the latest releases of 3 main libraries: JEFF-3.2, JENDL-4.0 and ENDF/B-VII.1, as shown for the 400 individuals benchmarks in fig. 2. This figure represents the status of the current knowledge on  $^{239}\text{Pu}$  criticality experiments. The  $C/E$  ratios (being the ratios of calculation over experiment) are presented for three sources of nuclear data: the American library ENDF/B-VII.1 [3], the Japanese library JENDL-4.0 [4] and the OECD/NEA library JEFF-3.2 [6]. The order of the benchmarks in this figure is different than usually presented: from the lowest  $C/E$  values to the highest for the ENDF/B-VII.1 library, therefore mixing types of benchmarks. As indicated by the  $\chi^2$  values, the agreement is not (yet) perfect, with the origin of the discrepancies lying in the  $^{239}\text{Pu}$  evaluation (for

instance imperfect cross sections), in the benchmark itself (incorrect estimation of the experiment), or within the evaluation of other isotopes included in the benchmark. Coming back to the ultimate goal of the nuclear data evaluators, a successful work is achieved if a  $^{239}\text{Pu}$  evaluation agrees with differential data (for instance pointwise cross sections) and if a large set of  $C/E$  for integral data as presented in fig. 2 are within the uncertainty bands (in other terms, if  $\chi^2$  values are as low as possible).

How to improve such agreements and to eventually get uncertainty due to nuclear data is the subject of this paper. In the following, we will present a method to ameliorate the  $C/E$  ratios by selecting specific  $^{239}\text{Pu}$  nuclear data files in association with the mentioned three libraries. The production of such evaluated data files is briefly presented in the first part of this paper. This first part is not new, and the production of random nuclear data was presented in a few dedicated papers (see, for instance, refs. [7, 8]). The use of these random files is presented in a second part. This is partially new, since a ranking of random files was already presented (see ref. [9]), but the new step described here concerns the combination of a few best files with a specific library (such as JEFF-3.2) to reach better  $C/E$  ratios and smaller uncertainties.

In summary, the method proposed in this paper is relatively simple and is based on two main points: 1) production of random nuclear data using a reaction code and rank them by comparison with integral measurements, and 2) provide a linear combination of random nuclear data files and weights (2 or more) to be used for simulations of nuclear systems. The advantages of this approach are double: 1) by using a combination of nuclear data files, the agreement with the a large set of benchmarks is greatly



**Fig. 3.** Various random fission cross sections for  $^{239}\text{Pu}$  (top) and uncertainties (bottom) in the resonance and fast neutron range. Also presented are the fission cross sections from the main libraries. For the uncertainties, the standard deviation from the three main libraries (ENDF/B-VII.1, JEFF-3.2 and JENDL-4.0) is used to indicate a dispersion, called “World library”.

improved and 2) uncertainties due to nuclear data can be calculated taking into account measurements and theoretical calculations (going from a conservative approach to best estimate plus uncertainties).

## 2 Random $^{239}\text{Pu}$ nuclear data files

The global search for best files is based on the idea of using random files with a quality flag and was extensively presented in ref. [9]. With the TALYS system [7], a number of 10000 ENDF random files for  $^{239}\text{Pu}$  are produced and processed into ACE files at room temperature with the NJOY processing code [10]. The use of random ENDF files was previously presented in many papers (see for instance refs. [11, 12] for the most recent ones). Figure 3 and table 1 present the most important cross sections and emitted particles for transport applications.

The selection of model parameters (and their uncertainties) is usually performed such that the calculated cross sections reproduce the experimental differential data. This type of work is performed by the nuclear data evaluation community, is time consuming but nevertheless necessary to obtain a sound evaluation (central values of our random cross sections). In this work, a slightly different approach is taken, by not directly fitting experimental data, but instead fitting a selected evaluation. This has at least two advantages as long as the evaluation of reference is trusted: 1) gain of time by not re-analyzing experimental data, and 2) directly adjusting model parameters to produce cross sections which will lead to acceptable integral calculations. The evaluation of reference in the following is from the JEFF-3.2 library since it has shown good performance.

As seen in table 1, the three main libraries significantly differ for some reactions, which can be translated into a spread (a standard deviation), indicating a degree of knowledge among many nuclear data evaluators. For the

goal of the present work, the spread of the random files should be based on the experimental and theoretical uncertainties, but should also allow for additional spread to cover a large set of the “nuclear phase space”.

## 3 Weights based on integral benchmarks

Integral benchmarks as defined in the ICSBEP database [1] are one of the most common tools used to assess the quality of a nuclear data library. A large set of cases exists for the most important isotopes (such as  $^{235,238}\text{U}$  and  $^{239}\text{Pu}$ ) and many of them are traditionally used (see for instance the benchmarking work presented in ref. [13]). In the case of  $^{239}\text{Pu}$ , more than 400 different benchmark cases can be found in ref. [1]. Clearly, in the goal of benchmarking a nuclear data library using these  $k_{\text{eff}}$  benchmarks, one should consider the largest number possible of benchmarks in order to obtain a meaningful statistical distribution of cases from which different characteristics can be extracted. In this process, two aspects always need to be carefully considered: 1) the quality, and thus reliability, of the integral experiment itself, and 2) the influence of other materials present. In practice, a compromise has to be found between the number of considered benchmarks and the required calculation time to obtain results.

### 3.1 Selection of benchmarks and libraries

In the case of benchmarking a single library, calculating 400 benchmarks is possible with a single computer (possibly done within a week). In the present case, as we want to rank each random  $^{239}\text{Pu}$  file as a function of the agreement with the benchmarks, it is necessary to calculate each benchmark ten thousand times, since ten thousands random  $^{239}\text{Pu}$  files are used. Additionally, as we are considering the three main world libraries, each set of ten thousand calculations for a single benchmark should be performed 3 times. Considering the limited computational power available, a subset of 35  $^{239}\text{Pu}$  benchmarks is selected, as presented in table 2. In this list, 22 benchmarks are classified as “fast” and 13 are “thermal”. This selection comes from the availability of their description as an input to the MCNP code and the intention of covering a large set of cases. These benchmarks contain a large amount of  $^{239}\text{Pu}$  and it is expected that the random  $^{239}\text{Pu}$  cross sections have a strong effect.

Figure 4 presents the current state of simulations for these benchmarks considering the three most accepted libraries: ENDF/B-VII.1, JENDL-4.0 and JEFF-3.2.

As seen in this figure, some libraries globally perform better than the others, depending on the amount of efforts dedicated by each evaluation team. In this study, we are using the  $\chi^2$  definition to measure the quality of a library, as presented in eq. (1), with  $n = 35$ . Within these 35 benchmarks, one has to use care when searching for the best  $\chi^2$ . For instance, the minimization of  $\chi^2$  as defined in eq. (1) will have the tendency to promote benchmarks with  $C/E$  far from 1, compared to benchmarks with  $C/E$

**Table 1.** Some nuclear quantities (cross sections, average energies) for  $^{239}\text{Pu}$  of the main libraries compared to the average from the 10000 random files. xs means cross sections, th means thermal energy,  $\bar{\nu}$  is the average prompt neutron fission emission,  $\bar{E}^{\text{PFNS}}$  means average energy for the prompt fission neutron spectra.

| Quantity  | JEFF-3.2 | JENDL-4.0         | ENDF/B-VII.1      | Random files |             |
|---|----------|-------------------|-------------------|--------------|-------------|
| $\text{xs}_{\text{th}}(\text{n},\text{f})$ (b.)     | 747.2    | $742.8 \pm 8.7$   | $747.1 \pm 8.7$   | 750.8        | $\pm 26.0$  |
| $\text{xs}_{\text{th}}(\text{n},\gamma)$ (b.)       | 270.1    | $275.8 \pm 4.7$   | $271.4 \pm 5.1$   | 271.8        | $\pm 0.3$   |
| $\text{xs}_{\text{th}}(\text{n},\text{el})$ (b.)    | 8.0      | $8.8 \pm 0.1$     | $8.8 \pm 0.1$     | 7.9          | $\pm 0.1$   |
| $\text{xs}_{2\text{MeV}}(\text{n},\text{inl})$ (b.) | 0.68     | 0.82              | 1.85              | 1.61         | $\pm 0.06$  |
| $\bar{\nu}_{\text{th}}$ (-)                         | 2.875    | $2.878 \pm 0.003$ | $2.879 \pm 0.005$ | 2.876        | $\pm 0.010$ |
| $\bar{E}_{\text{th}}^{\text{PFNS}}$ (MeV)           | 2.11     | 2.11              | 2.11              | 2.11         | $\pm 0.10$  |
| $\bar{E}_{0.5\text{MeV}}^{\text{PFNS}}$ (MeV)       | 2.13     | 2.12              | 2.12              | 2.08         | $\pm 0.10$  |
| $\bar{E}_{1\text{MeV}}^{\text{PFNS}}$ (MeV)         | 2.14     | 2.12              | 2.14              | 2.11         | $\pm 0.10$  |

**Table 2.** List of the  $^{239}\text{Pu}$  benchmarks selected. Each criticality benchmark name consists of three letters: the first one defining the type of enrichment (p for a plutonium benchmark), the second one defining the physical form (m for metal, s for solution) and the last one defining the neutron spectrum (f for fast, t for thermal and i for intermediate).

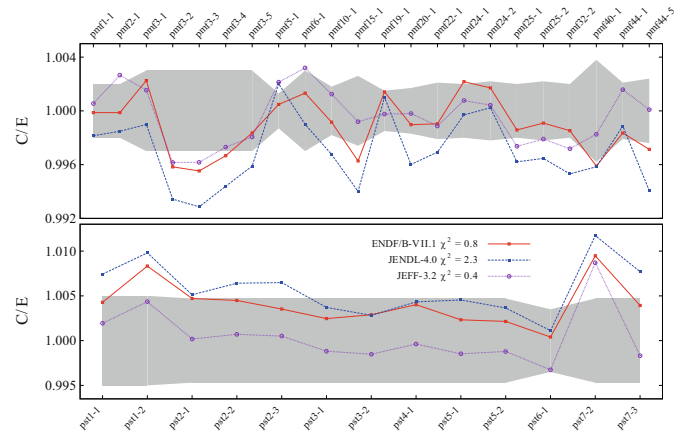
| Fast benchmarks    |         |         |         |         |
|--------------------|---------|---------|---------|---------|
| pmf1-1             | pmf2-1  | pmf3-1  | pmf3-2  | pmf3-3  |
| pmf3-4             | pmf3-5  | pmf5-1  | pmf6-1  | pmf9-1  |
| pmf10-1            | pmf15-1 | pmf19-1 | pmf20-1 | pmf22-1 |
| pmf24-1            | pmf24-2 | pmf25-1 | pmf25-1 | pmf32-2 |
| pmf40-1            | pmf44-1 | pmf44-5 |         |         |
| Thermal benchmarks |         |         |         |         |
| pst1-1             | pst1-2  | pst2-1  | pst2-2  | pst2-3  |
| pst3-1             | pst3-2  | pst4-1  | pst5-1  | pst5-2  |
| pst6-1             | pst7-2  | pst7-3  |         |         |

relatively close to 1, simply because the gain for the *deviant* benchmarks is greater than for the others. Using eq. (1), the ENDF/B-VII.1 library obtains a value of 0.8, JEFF-3.2 0.4 and JENDL-4.0 2.3. In the case of calculations leading to all  $C/E$ s equal to 1 with a certain library, the  $\chi^2$  value would be 0.

From fig. 4, one can observe that all libraries provide  $C/E$  which are outside the benchmark uncertainties for some benchmarks: pmf3-2, pmf3-3 and pst7-2. This can have three reasons: 1) the nuclear data (for  $^{239}\text{Pu}$  or other isotopes) are not compatible with the benchmark, 2) the benchmark is not well evaluated, and 3) a mixture of the two first reasons.

As explained, each of the 35 benchmarks needs to be calculated 30000 times, once for each random file. For instance, the pmf1 benchmark is calculated 10000 times for each library (considering 3 libraries, the number of calculation is  $3 \times 10000$ ).

Based on this selection of 35 benchmarks, each random  $^{239}\text{Pu}$  file can be used in combination with another library (for instance ENDF/B-VII.1, of JEFF-3.2 for the



**Fig. 4.**  $C/E$  for the 35 considered criticality benchmarks for the three main libraries ENDF/B-VII.1, JENDL-4.0 and JEFF-3.2.

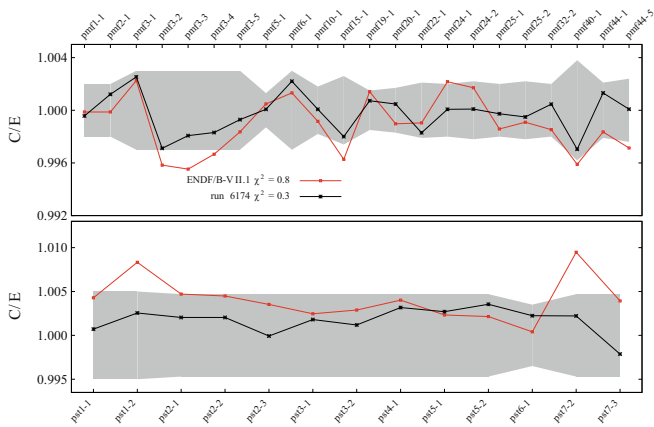
other isotopes such as  $^{240}\text{Pu}$ ,  $^{56}\text{Fe}$ , etc.) to calculate 35 benchmark values (or  $C/E$ ). Therefore, for each random  $^{239}\text{Pu}$  file, 35  $C/E$  values are obtained and can be used in eq. (1) to calculate a single  $\chi^2$  value per random file. Nonetheless, this distribution (for each library) of 10000  $C/E$  for each of the 35 benchmarks can be used to extract the “best” possible solution, as presented in the following sections. The definition of “best” is not univocal, and 3 solutions will be presented.

### 3.2 A good fit

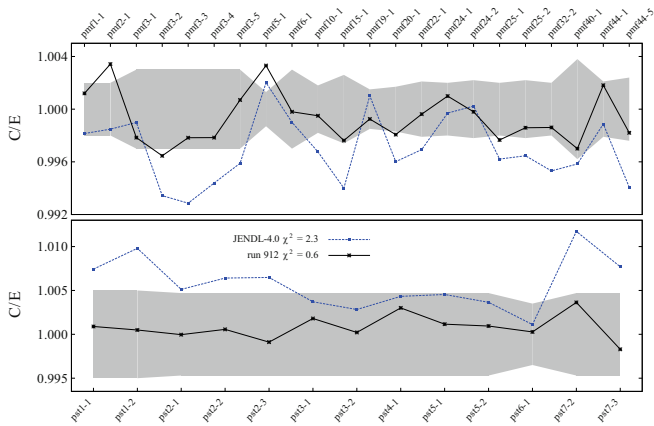
As mentioned in the previous paragraph, each random  $^{239}\text{Pu}$  nuclear data can be associated with a  $\chi^2$  value as defined in eq. (1) for the given 35 benchmarks. These  $\chi^2$  values are also dependent on the library associated with these random files (the same random file can give different  $C/E$  values if associated with JEFF-3.2 or JENDL-4.0). In refs. [9,14], the same exercise was performed with a unique library (ENDF/B-VII.0 in [9] and JEFF-3.1 in [14], but the selected benchmarks were different as well as the random isotope). From these  $\chi^2$  distributions, the best random files when associated with ENDF/B-VII.1, JENDL-

**Table 3.**  $\chi^2$  values using the three best files in combination with each library, considering 35 benchmarks. The last column presents the  $\chi^2$  for a given random file, averaged over the three libraries.

| File   | ENDF/B-VII.1 | JEFF-3.2    | JENDL-4.0   | Average $\chi^2$ |
|--------|--------------|-------------|-------------|------------------|
| 6174   | <b>0.26</b>  | 0.49        | 0.69        | 0.48             |
| 5989   | 0.33         | <b>0.38</b> | 0.72        | 0.48             |
| 912    | 0.34         | 0.45        | <b>0.60</b> | 0.46             |
| No     |              |             |             |                  |
| random | 0.83         | 0.64        | 2.30        | 1.26             |



**Fig. 5.**  $C/E$  for ENDF/B-VII.1 compared to the combination ENDF/B-VII.1-run 6174.

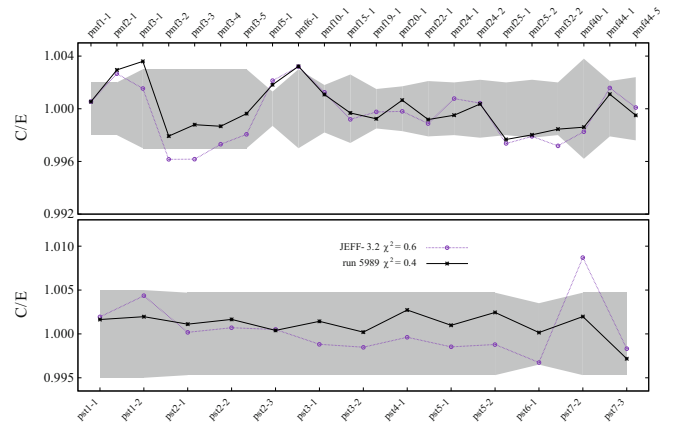


**Fig. 6.**  $C/E$  for JENDL-4.0 compared to the combination JENDL-4.0-run 912.

4.0, or JEFF-3.2 can be extracted (with the lowest  $\chi^2$ ) and are presented in table 3. In order to see the gain for each of the  $C/E$  values, figs. 5 to 7 present the comparisons of  $C/E$  for three libraries with their best random files.

In each case the gain in  $\chi^2$  value can be seen at the  $C/E$  values as  $\chi^2$  values are strongly reduced: for ENDF/B-VII.1, the majority of the gain is for the thermal benchmarks, for JENDL-4.0 both fast and thermal benchmarks are improved and for JEFF-3.2 the gain is primarily for a limited set of benchmarks.

An important remark on the values of the presented  $\chi^2$  should be noticed. It is correct that the values of the  $\chi^2$  are



**Fig. 7.**  $C/E$  for JEFF-3.2 compared to the combination JEFF-3.2-run 5989.

strongly reduced, but the original  $\chi^2$  for two of the three libraries were already lower than 1, indicating that any improvement would not be statistically significant. These examples are nevertheless indicative for the understanding of the method. The values of  $\chi^2$  smaller than 1 come from the restricted selection of benchmarks. As presented later in the paper, the  $\chi^2$  values for the three main libraries degrade when more benchmarks are considered (getting larger than 1). Therefore, the results presented in this section present the advantage to be illustrative, rather than leading to a final improvement of the evaluations.

The three best files for each library (ENDF/B-VII.1, JENDL-4.0 and JENDL-4.0) are different, but each of this best file also lead to an acceptable  $\chi^2$  if associated with a different library. To illustrate this remark, table 3 presents the different  $\chi^2$  values for the best files with their associated library and with different libraries. As seen, file 6174 is the best random file associated with ENDF/B-VII.1, but it also gives good  $\chi^2$  values with the other libraries (it is the 7th best file for JENDL-4.0 and the 4th best for JEFF-3.2). Similarly, good results are obtained with file 912 with JEFF-3.2 and ENDF/B-VII.1, and also with file 5989. Therefore, any of these files gives an acceptable result independently of the selected original library (ENDF/B-VII.1, JEFF-3.2 or JENDL-4.0), and generally better than the original library alone.

As presented in the following, lower  $\chi^2$  can be obtained if combination of random files with specific weights are used instead of a single random file.

### 3.3 A better fit: Combination of distributions

In the following are presented the main hypothesis and results obtained by applying this simple classification method.

- 35 criticality benchmarks are considered;
- 10000 random  $^{239}\text{Pu}$  nuclear data files are used;
- each random file is associated with its  $\chi^2$  weight as defined in eq. (1);
- random file 6174 has the smallest  $\chi^2$  with the ENDF/B-VII.1 library;

**Table 4.**  $\chi^2$  values for 35 benchmarks in different cases (a unique library, or a library coupled with a selection of random  $^{239}\text{Pu}$  files.

|   | Case (1) | Case (2) | Case (3)    | Library + best file | Library only |
|---|----------|----------|-------------|---------------------|--------------|
| ENDF/B-VII.1                            | 0.6      | 0.4      | 0.2         | 0.3                 | 0.8          |
| JENDL-4.0                               | 1.1      | 0.7      | 0.5         | 0.6                 | 2.3          |
| JEFF-3.2                                | 0.8      | 0.5      | 0.3         | 0.4                 | 0.6          |
| Number of calculations for each library | $10^4$   | $10^4$   | $\simeq 10$ | 1                   | 1            |

- random file 912 has the smallest  $\chi^2$  with the JENDL-4.0 library;
- random file 5989 has the smallest  $\chi^2$  with the JEFF-3.2 library.

Up to now, we have succeeded to find a random set of nuclear data for  $^{239}\text{Pu}$  which will improve the agreement between a given library (where the original  $^{239}\text{Pu}$  data are exchanged with the random ones) and the set of defined benchmarks.

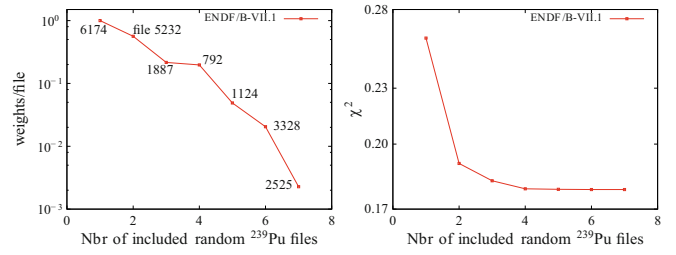
It is also possible to consider a limited number of random files (with a minimum of 2) and to combine them in order to improve the global agreement to benchmarks (smaller  $\chi^2$ ). In the case of using a combination of a few random files, a central value  $\overline{C}_i$  can be obtained for each benchmark  $i$ , summing the results for each random file  $j$ :

$$\overline{C}_i = \frac{1}{\sum w_j} \sum w_j \times k_{\text{eff},j}. \quad (2)$$

A few possibilities on how to combine random files exist among which

- 1) take all random files with an equal weight of 1 (similar to the Total Monte Carlo (TMC) approach [8]);
- 2) take all random files with their weights being equal to  $\exp(-\chi_j^2)$  ( $j$  is the random file index), and
- 3) take a subset of  $m < n$  random files (using  $m$  files having the best  $\chi^2$ , and  $n$  being the number of benchmarks) and search for specific positive weights to lower the  $\chi^2$  obtained from the weighted combination.

All these possibilities rely on combination of results, implying that an eventual user would have to repeat a given calculation a few times (with different nuclear data files). This additional burden can eventually be justified if the combined  $C/E$  show substantial improvement, or if an additional quantity (such as the uncertainty due to nuclear data) can be obtained. As an example of the gain obtained in these different 3 cases, table 4 presents the  $\chi^2$  for the 35 previous benchmarks using eq. (1), but the  $C_i$  values are changed by  $\overline{C}_i$  as defined in eq. (2).  $w_j$  are the weights of the considered random files and  $C_i$  are the calculated  $k_{\text{eff}}$  for the benchmark  $i$  using the considered random files (in case (1),  $w_j = 1$ , in case (2),  $w_j = \exp(-\chi_j^2)$  and in



**Fig. 8.** Left: weights of the added files in the search to minimize  $\chi^2$  by using a weighted sum of solutions for 35 benchmarks. Weights are presented for ENDF/B-VII.1. Numbers indicate the random file number. Right: Obtained  $\chi^2$  by using a weighted combination of results (log scale).

case (3),  $w_j$  are obtained from a minimalization of the  $\chi^2$  for the 35 benchmarks).

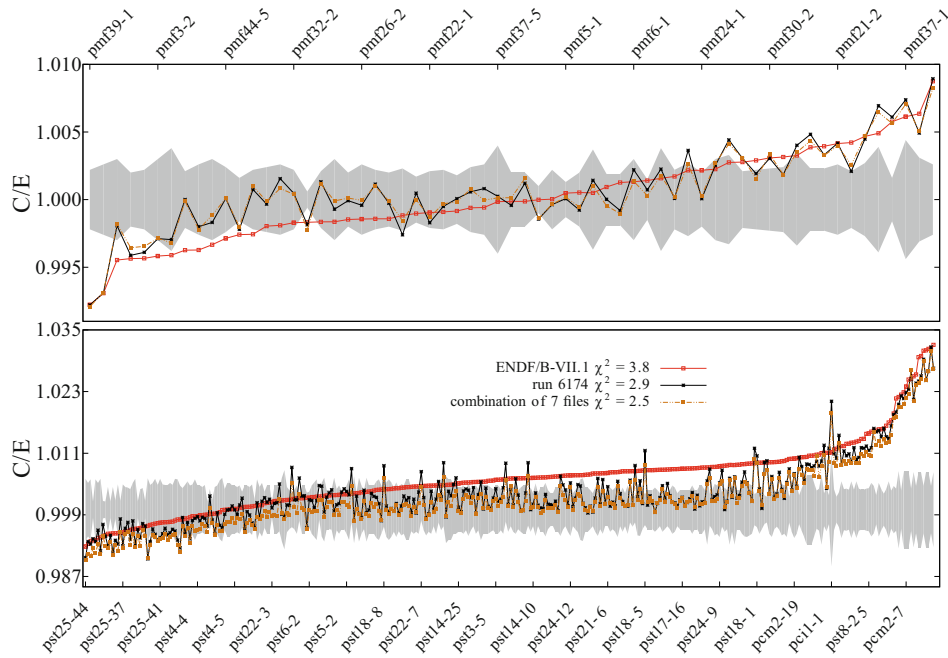
The solutions (1) and (2) are not practical since they involve too many calculations, for no gain, compared to the  $\chi^2$  obtained with the best random file. Finally, only case (3) can be of interest considering these 35 benchmarks.

The search for the random files and their associated weights for the case (3) can be performed with different algorithms. In the case of the solution (3), the specific weights and the number of files can be searched with a random process, such as iteratively adding one file and randomly searching its weight to minimize the combined  $\chi^2$  (if no weight can be found leading to a better  $\chi^2$ , the file is discarded). An example of such combination of files and weights is presented in fig. 8 for the case (3).

In the case (3), the  $\chi^2$  as a function of included random files are indicating a lower limit, reached with a relatively low number of random files. Furthermore, the weights of the added random files are decreasing and addition of new random files will not lower  $\chi^2$ . Therefore, with the case (3), a limited number of random files (and their weights) are enough to achieve a low and stable  $\chi^2$ : 7 random files for ENDF/B-VII.1, 6 for JENDL-4.0 and 9 for JEFF-3.2.

It can be noticed that the  $\chi^2$  presented in table 4 are almost all less than 1, meaning that the obtained improvements for  $C/E$  are within the experimental standard deviations. In the case of a larger study, including benchmarks with  $C/E$  far from 1, and adjusting more isotopes at a time (including structural materials and other actinides), the updated  $\chi^2$  are expected to be greater than 1, as shown in the next section.

As noticed, different  $\chi^2$  values are obtained depending on the choice of the original library. For instance, the use of the ENDF/B-VII.1 library leads to the smallest  $\chi^2$ , whereas JENDL-4.0 leads to the highest one. The origin of the difference of  $\chi^2$  lies in the isotopes other than  $^{239}\text{Pu}$ . The presented method proposes modifications for  $^{239}\text{Pu}$ , thus compensating for possible deficiencies in the  $^{239}\text{Pu}$  evaluations, but also in the evaluations of other isotopes, such as  $^{241}\text{Pu}$ . There is therefore a limitation in the possible improvement, when changing one isotope. If the present method would be applied to all the important isotopes for



**Fig. 9.**  $C/E$  for a large number of  $^{239}\text{Pu}$  benchmarks using the ENDF/B-VII.1 library, ENDF/B-VII.1 + run 6174 and ENDF/B-VII.1 with a combination of 7 random files.

the considered benchmarks, one can imagine that all  $\chi^2$  for the three (updated) libraries would be equivalent.

### 3.4 Predictive power for $k_{\text{eff}}$

In the previous sections, we have seen different ways to improve  $C/E$  results based on existing libraries and a selection of nuclear data. The first one, using the random run which provides the best  $\chi^2$  instead of the file delivered with the original library (ENDF/B-VII.1, JEFF-3.2 or JENDL-4.0) has shown good performances for the selected benchmarks. Another solution, using a weighted distribution of well-selected random files has also shown good results.

It is relatively simple to test the prediction power of these two methods, simply by calculating  $C/E$  for additional benchmarks. For this, we have calculated a total of 400  $^{239}\text{Pu}$  criticality benchmarks (as presented in fig. 2) for the main three libraries and their combined best files. Results for the ENDF/B-VII.1 library are presented in fig. 9.

Similar plots were obtained for JENDL-4.0 and JEFF-3.2 but are not reproduced here. Considering this set of 400 benchmarks, the global  $\chi^2$  values for each library with its best random files are presented in table 5. The  $\chi^2$  values for each libraries with  $n$  random files are indeed lower compared to the  $\chi^2$  values using only one single best file, but the differences are not considerable. From the point of view of a library performance regarding criticality benchmarks, the gain by using many best files is not as justified as in the case of using one single best file, taking into account the additional required calculation time. As presented in the next section, the advantage of using  $n$

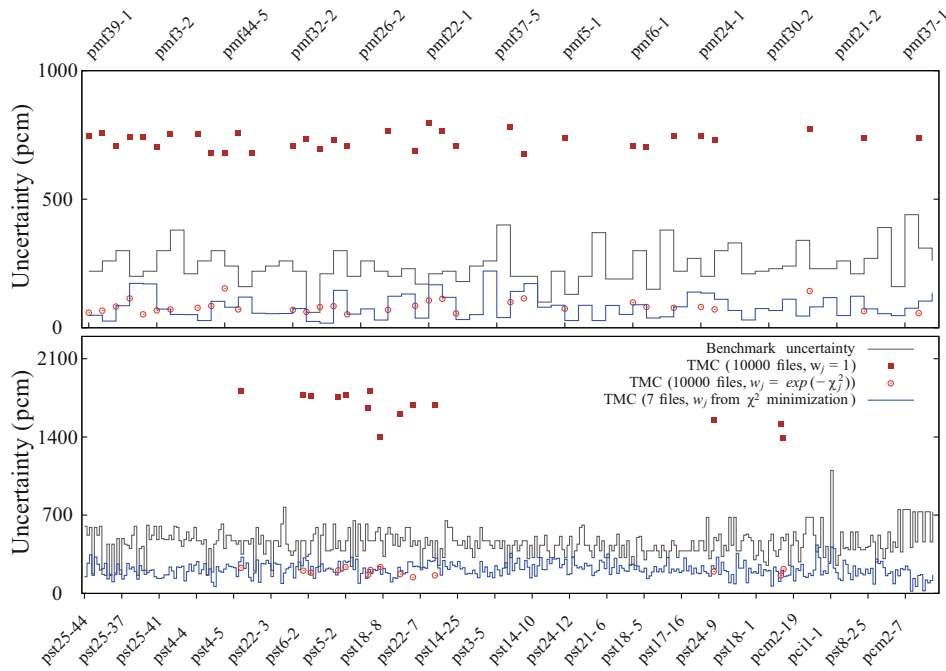
**Table 5.**  $\chi^2$  values for 400 benchmarks in different cases (the original library, or coupled with a selection of random  $^{239}\text{Pu}$  files).

| Library      | Case         |                      |                          | $n$ |
|--------------|--------------|----------------------|--------------------------|-----|
|              | Library only | Library +1 best file | Library + $n$ best files |     |
| ENDF/B-VII.1 | 3.8          | 2.9                  | 2.5                      | 7   |
| JENDL-4.0    | 4.9          | 2.6                  | 2.4                      | 6   |
| JEFF-3.2     | 3.2          | 2.5                  | 2.4                      | 9   |

(well-chosen) random files lies in the possibility of obtaining uncertainties.

## 4 Reduction of uncertainties

As presented in the previous sections, the method of selecting a set of random files allows to obtain better  $C/E$  values. These selected files come with specific weights, and the  $k_{\text{eff}}$  results of the simulations using these files and weights lead to a specific distribution. A similar approach is used in the Total Monte Carlo method (TMC), where hundreds (or more) calculations are repeated with all available random files. The TMC method with equal weights for all random files implies that all files are used for the uncertainty quantifications, independently of their quality regarding benchmark performances. Of course all the random files were first produced taking into account the available differential data (as in the present case). It can be argued that using equal weights leads to a conservative estimation of the uncertainties, as it does not take



**Fig. 10.**  $C/E$  uncertainties for 400  $^{239}\text{Pu}$  benchmarks using the TMC approach, and the uncertainty reduction method using eq. (3).

into account all possible sources of information (such as the integral benchmarks). In the present case, all the random files are ranked regarding their integral performances and given a specific weight. Therefore not only the differential, but also the integral information is included in the selection, and the use of these specific random files reflects a better estimation of the nuclear data uncertainties compared to a conservative approach.

In practice, the uncertainties are obtained as in the case of the TMC method: repeating the same calculation  $n$  times, each time changing the input nuclear data files.  $n$  is, as presented in table 5 a small number. From the  $n$  calculations,  $n$  results are obtained and a standard deviation can be extracted. As a weighted average was used to obtain a calculated value (see eq. (2)), the standard deviation  $\sigma$  is:

$$\sigma_i^2 = \frac{1}{\sum w_j} \sum [w_j \times (C_j - \bar{C}_j)^2]. \quad (3)$$

With a selection of 7 files and their associated weights in the case of the ENDF/B-VII.1 library, eq. (2) leads to the uncertainties presented in fig. 10 (again, similar figures can be obtained for the JEFF-3.2 and JENDL-4.0 library, using either 6 or 9 random files).

One can see that in almost all cases an important decrease of the uncertainties compared to the TMC method with  $w_j = 1$ . An alternative way of calculating the uncertainties is to consider all the random files, but this time weighted by their  $\exp(-\chi_j^2/2)$  values. This leads to results very close to the ones using only 7 files with their associated weights coming from the minimization of the  $\chi^2$ . The drawback of this last method is the necessity of running a simulation a large number of times, whereas with 7 selected random files, one can obtain uncertainties in 7 runs.

In other words, the current procedure leads to a limited number of random files, each with their own weight, which leads to the same calculated uncertainty as obtained with many random files.

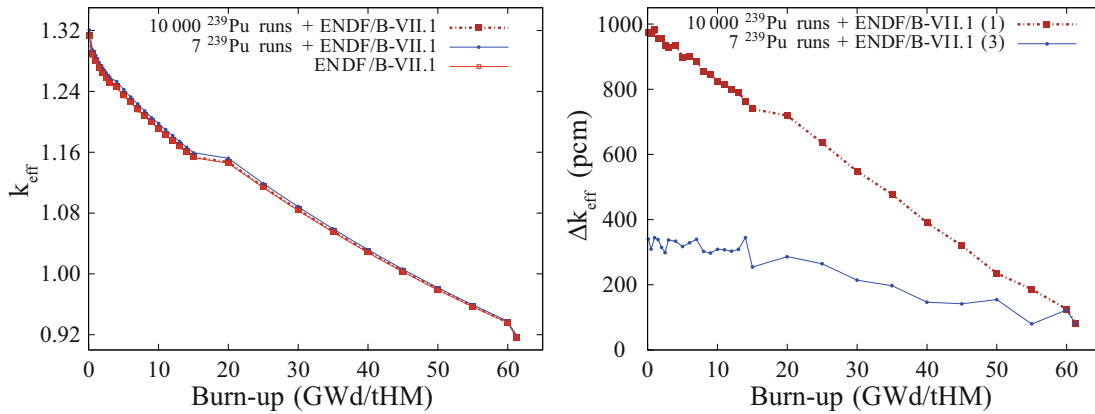
The usage of eq. (2) to calculate the central value of a series of simulations and of eq. (3) to extract a standard deviation assures for consistency between these two quantities. It is unfortunately a current practice to use different simulation sets or sources to calculate a quantity (for instance with the ENDF/B-VII.1 library) and also to provide uncertainties (with the covariances coming from SCALE package).

An additional example for the uncertainty reduction is presented in fig. 11, where a simple typical MOX pincell model as described in ref. [15] is used with the SERPENT code [16] (the  $^{239}\text{Pu}$  content is about 4.3%).

As in the case of the criticality benchmarks, the uncertainties are strongly reduced when using the selection of random files with their weights. This reduction represents the propagation of information from the criticality benchmarks (used to select files and weights) to the given application (here a simple pincell model). In terms of methodology, the higher uncertainties (without taking into account the information from integral benchmarks such as the criticality benchmarks) can be considered as *conservative*, compared to the lower uncertainties (taking into account the information from integral benchmarks), which are in turn a *better estimate*.

As a side remark, the value of  $n$  depends on the used library and on the selection of benchmarks. Similar comments can be applied to the selected random files and their weights. It is therefore important to use the selected random files with their associated original library, otherwise the whole consistency of the method is not guaranteed. Be-





**Fig. 11.** Burn-up calculations for in MOX pincell:  $k_{\text{eff}}$  (left) and  $\Delta k_{\text{eff}}$  (right). The  $^{239}\text{Pu}$  nuclear data are varied: without uncertainty reduction (10000 calculations,  $w_j = 1$ , case (1)) and with uncertainty reduction (7 weighted calculations, case (3)).

cause of our restricted computational power, we could not apply this method using more benchmarks, or for instance to other isotopes. But more sizable results can be obtained if applied to isotopes such as  $^{238}\text{U}$  and  $^{235}\text{U}$ , using a large set of benchmarks (about 2000) for the selection. Furthermore, not only criticality benchmarks can be used, but any type of reliable simulation including experimental results (such as data from reactor surveillance programs).

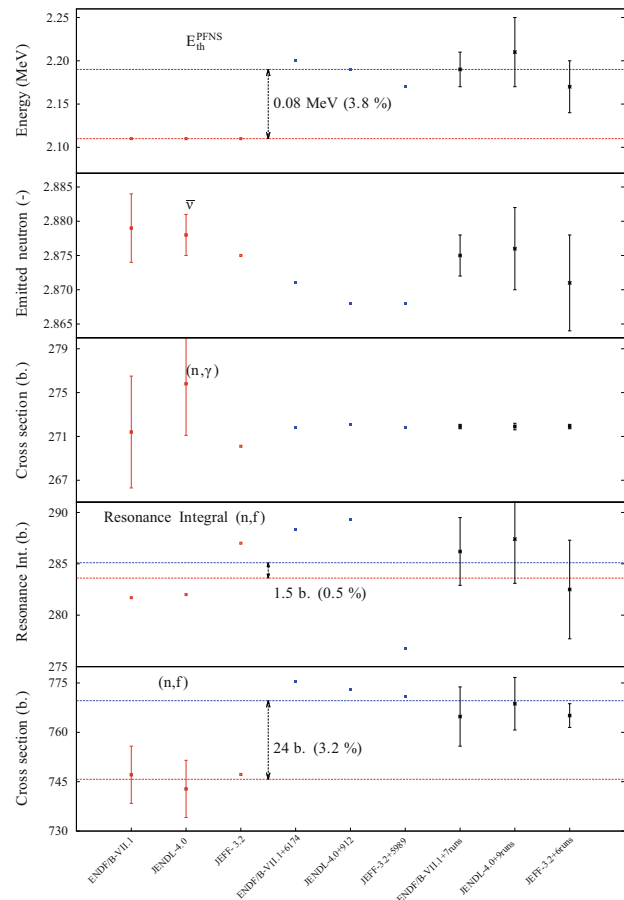
### 5 Feedback to nuclear data evaluation

Different information on the differential nuclear data for  $^{239}\text{Pu}$  can be extracted from the present study. The better agreement with the integral benchmarks obtained with the different random files and their combinations (compared to the original libraries) can be interpreted in terms of differential data. These remarks are valid under the assumptions that these better agreements are obtained thanks to an improvement of the differential data, and not to compensation effects, for instance between the  $^{239}\text{Pu}$  data and some other (wrongly assessed) isotope data used in the considered set of benchmarks. This assumption is of course important, and the conclusions taken in this section will then be indicative rather than definitive.

As we are also using the information from the results coming from the combination of random files (for instance from the JEFF-3.2 library coupled with a linear combination of 6 random files), the presented conclusions are also valid under the assumptions that a linear combination of ACE files (or MCNP runs) can be translated into a linear combination of ENDF files (or differential nuclear data). This second assumption is also subject to caution, since a  $k_{\text{eff}}$  value is not a linear combination of cross sections. Therefore the updated differential nuclear data given in the following should be understood as indications rather than strong recommendations.

Figure 13 is presenting the ratios of cross sections in the case of the  $^{239}\text{Pu}$  nuclear data from the ENDF/B-VII.1 library and from run 6174; fig. 12 is summarizing the main characteristics of the libraries and their associated random files.

As seen in fig. 13 local changes to cross sections can be relatively important (up to 20% in the case of the fission

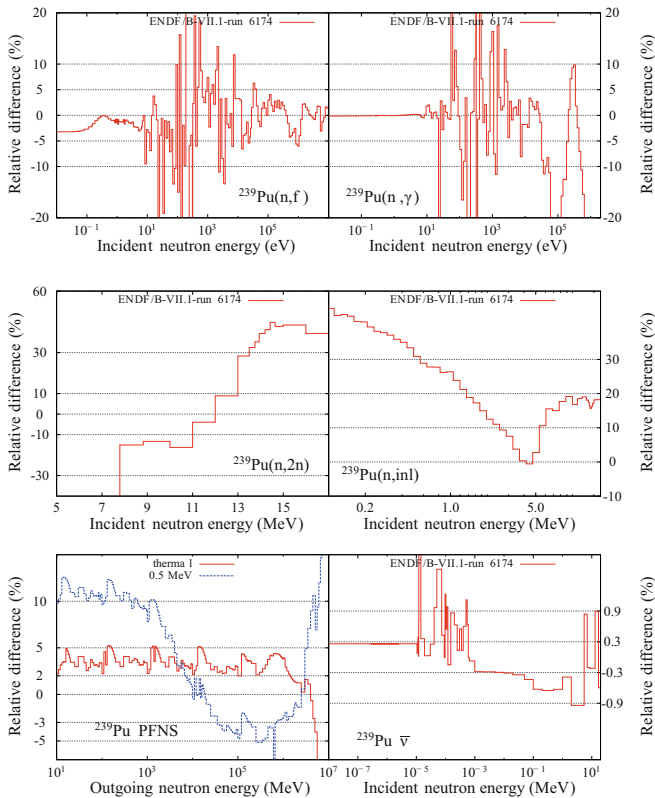


**Fig. 12.** Comparison of some important differential nuclear data between the main three libraries and their associated random files. The dotted lines are linear fits using either the three main libraries only (in red), or the data from this study (in blue).

and capture cross sections), but more importantly are the specific changes in some restricted energy regions and for some cross sections. Figure 12 is presenting the changes of cross sections and other quantities in the thermal range.

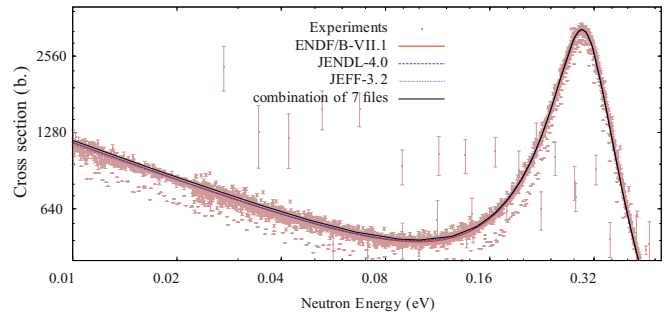
The following remarks can be mentioned:

- 1) Thermal capture cross section. No sensible changes to the thermal capture cross section are noticed.



**Fig. 13.** Comparison between the ENDF/B-VII.1  $^{239}\text{Pu}$  nuclear data and the best random file associated with ENDF/B-VII.1: run 6174. As the comparison involves a large set of numbers, results are presented only for a the main quantities.

- 2) Thermal prompt nu-bar ( $\bar{\nu}$ ). As for the thermal capture cross section, the present study does not indicate changes in  $\bar{\nu}$  at thermal energy. There is however a tendency to slightly lower  $\bar{\nu}$  above 1 keV.
- 3) Average energy of the prompt fission neutron spectrum (PFNS) at thermal incident energy. As indicated in fig. 12, the data from the three single random files (run 6174, 912 and 5989) associated with the main three libraries indicate an increase of the average energy of the PFNS at thermal incident energy. The increase is in the order of 80 keV, making the thermal spectrum a bit “harder”. For the 500 keV incident energy, the average energy is decreased by about 700 keV, making it a bit “softer” (from 2.12 MeV for the three main libraries to 2.05 MeV for the random files). This study cannot nevertheless assess the impact of the PFNS only.
- 4) Thermal fission cross section. As seen in fig. 12, there is an indication that better integral results are obtained by increasing the thermal fission cross section by about 3.2%. This change is relative high for such an important cross section, but it should be noticed, as presented in fig. 13, that close to the important resonance at 0.29 eV, the cross section is almost not changed. This remark can be extended to the fission resonance integral (see fig. 12) which is changed by a relatively small value of 0.5%. This is also illustrated in fig. 14



**Fig. 14.** Fission cross sections in the thermal range for the main libraries, measured data and the combination of 7 files.

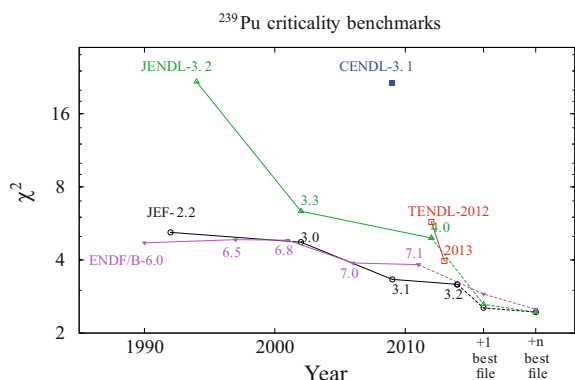
where the pointwise fission cross section up to 0.5 eV from the different libraries are compared.

- 5) Inelastic cross section. As presented in fig. 13, the total inelastic cross section is significantly reduced compared to the ENDF/B-VII.1 original library (as between JENDL-4.0 and run 912). In the case of the JEFF-3.2 library, the inelastic cross section of run 5989 is slightly reduced. It is difficult to judge from this study the impact of the inelastic cross section, but the experimental uncertainties for such cross section are certainly large enough to allow for some important changes.

As a general comment on the presented method and examples, this work is by itself not enough to produce a general purpose evaluation. It is here focused on the impact on criticality benchmarks, without further analysis of other applications. Therefore the adjusted evaluation is only one step towards a complete and general nuclear data evaluation which can be used in any type of application. But it is relatively straightforward to generalize the presented approach to other types of benchmarks (for instance decay heat or fuel storage).

Additionally, as for any quality evaluation, more efforts should be given in the comparison between the proposed updated evaluation and the experimental data. One can imagine a similar procedure as presented here, not with  $k_{\text{eff}}$  values but with a selection of experimental data. The presented work outlines general principles and ideas. It is not presenting a final ( $^{239}\text{Pu}$ ) evaluation, but rather a new method of analyzing integral data, which should be combined with more traditional evaluation work.

Finally, a missing information in this work is the sensitivity vectors. These vectors indicate if a cross section is sensitive to the selected benchmarks. If yes, then the conclusions for this specific cross section can be considered as significant. If not, the observed cross section variations can be understood as a result of a random process. As in any Monte Carlo approach, there is an intrinsic difficulty to obtain sensitivity vectors. But recent works are showing new possibilities as in ref. [17] (where a selected sampling can help to obtain sensitivity vectors), or in ref. [18] (where new methods of analysing data allow different ranking of importance).



**Fig. 15.** Same as fig. 1, but with the additional  $\chi^2$  obtained from the use of the best random files.

## 6 Conclusion

As explained in the introduction, the goal of a nuclear data evaluator is to provide recommended values based on differential and integral information. It was shown in fig. 1 that this work was successfully performed over the past decades, with a certain limitation in the recent years. To help the nuclear data evaluator providing better results, the present method based on the use of random nuclear data is proposed. As presented, this method allows to obtain better global agreement with integral data such as criticality benchmarks, while keeping the consistency with differential data (such as cross sections). As shown in fig. 15, the use of the present method significantly improve  $C/E$  ratios by selecting  $^{239}\text{Pu}$  cross sections. One could furthermore apply this method to other isotopes contained in the selection of benchmarks to improve even more  $C/E$  ratios: the full power of such method can be used for  $^{235}\text{U}$ ,  $^{238}\text{U}$ ,  $^{16}\text{O}$  and other important isotopes, combined with the use of a large set of integral benchmarks (criticality, shielding, activation, etc.). It should also be noticed that this approach can be combined with more traditional methods of nuclear data evaluation and would therefore become one of the many sources of observations that evaluators have to consider to produce a general purpose library. A second outcome of this selection of random files is the different estimation of uncertainties due to nuclear data. By choosing a set of random files with their weights based on  $C/E$  ratios for criticality benchmarks, this information can be used for neutron transport simulations.

Using this pre-selected set of random files is truly associating differential and integral knowledge into a separated specific neutronics application, therefore providing lower uncertainties than a conservative method.

## References

1. J.B. Briggs (Editor), *International Handbook of evaluated Criticality Safety Benchmark Experiments*, NEA/NSC/DOC(95)03/I, Organisation for Economic Co-operation and Development, Nuclear Energy Agency (2004).
2. J.F. Breisemeister, *MCNP - A General monte Carlo N-Particle Transport Code, Version 4C*, Los Alamos National Laboratory, Report LA-13709-M, April (2000).
3. M. Chadwick *et al.*, Nucl. Data Sheets **112**, 2887 (2011).
4. K. Shibata *et al.*, J. Nucl. Sci. Technol. **48**, 1 (2011).
5. Z.G. Ge, Y.X. Zhuang, T.J. Liu, J.S. Zhang, H.C. Wu, Z.X. Zhao, H.H. Xia, J. Kor. Phys. Soc. **59**, 1052 (2011).
6. The JEFF-3.2 library, available at the NEA Databank, [www.oecd-nea.org/dbdata/jeff/](http://www.oecd-nea.org/dbdata/jeff/).
7. A.J. Koning, D. Rochman, Nucl. Data Sheets **113**, 2841 (2012).
8. A.J. Koning, D. Rochman, Ann. Nucl. Energy **35**, 2024 (2008).
9. D. Rochman, A.J. Koning, Nucl. Sci. Eng. **169**, 68 (2011).
10. R.E. MacFarlane, A.C. Kahler, Nucl. Data Sheets **111**, 2739 (2010).
11. D.F. DaCruz, D. Rochman, A.J. Koning, Nucl. Technol. **185**, 174 (2014).
12. H. Sjostrand, E. Alhassan, J. Duan, C. Gustavsson, A.J. Koning, S. Pomp, D. Rochman, M. Osterlund, Nucl. Data Sheets **118**, 527 (2014).
13. S.C. van der Marck, Nucl. Data Sheets **113**, 2935 (2012).
14. D. Rochman, A.J. Koning, Nucl. Sci. Eng. **170**, 265 (2012).
15. P. Helgesson, D. Rochman, H. Sjostrand, E. Alhassan, A.J. Koning, Nucl. Sci. Eng. **177**, 321 (2014).
16. J. Leppanen, *PSG2/Serpent - A Continuous-energy Monte Carlo Reactor Physics Burnup Calculation Code*, VTT Technical Research Centre of Finland, Finland, <http://montecarlo.vtt.fi> (2010).
17. G. Shiba, Y. Kawamoto, M. Tsuji, T. Narabayashi, Ann. Nucl. Energy **75**, 395 (2015).
18. T. Zhu, A. Vasiliev, F. Ferroukhi, A. Pautz, *NUSS-RF: Stochastic Sampling-Based Tool for Nuclear Data Sensitivity and Uncertainty Quantification*, in *Proceeding of the International Conference Reactor Physics, PHYSOR2014, Kyoto, Japan, September 28, October 3, 2014*.