Benchmarking and validation activities within JEFF project

O. Cabellos$^{1,4}$, F. Alvarez-Velarde$^2$, M. Angelone$^3$, C.J. Diez$^1$, J. Dyrda$^1$, L. Fiorito$^4$, U. Fischer$^5$, M. Fleming$^6$, W. Häeck$^7$, I. Hill$^1$, R. Ichou$^7$, D. H. Kim$^8$, A. Klix$^5$, I. Kodeli$^9$, P. Leconte$^{10,1}$, F. Michel-Sendis$^4$, E. Nuppenmann$^5$, M. Pecchia$^{11}$, Y. Penelius$^{10,1}$, A. Plompen$^{12}$, D. Rochman$^{11}$, P. Ronojaro$^{13}$, A. Stankovskiy$^4$, J.Ch. Sublet$^6$, P. Tamagno$^{10}$, and S. van der Marck$^{14}$

$^1$ NEA, OECD, Paris, France
$^2$ CIEMAT, Madrid, Spain
$^3$ ENEA, Frascati, Italy
$^4$ SCK·CEN, Mol, Belgium
$^5$ KIT, Eggenstein-Leopoldshafen, Germany
$^6$ UKAEA, Abingdon, Oxfordshire, UK
$^7$ IRSN, Paris, France
$^8$ KAERI, Daejeon, South Korea
$^9$ JSI, Ljubljana, Slovenia
$^{10}$ CEA, DEN, DER, SPRC, Cadarache, France
$^{11}$ PSI, Villigen, Switzerland
$^{12}$ JRC, Geel, Belgium
$^{13}$ UPM, Madrid, Spain
$^{14}$ NRG, Petten, The Netherlands

Abstract. The challenge for any nuclear data evaluation project is to periodically release a revised, fully consistent and complete library, with all needed data and covariances, and ensure that it is robust and reliable for a variety of applications. Within an evaluation effort, benchmarking activities play an important role in validating proposed libraries. The Joint Evaluated Fission and Fusion (JEFF) Project aims to provide such a nuclear data library, and thus, requires a coherent and efficient benchmarking process. The aim of this paper is to present the activities carried out by the new JEFF Benchmarking and Validation Working Group, and to describe the role of the NEA Data Bank in this context. The paper will also review the status of preliminary benchmarking for the next JEFF-3.3 candidate cross-section files.

1. Introduction

The current evaluations of neutron induced cross sections are based on differential and integral experiments as well as nuclear reaction models. These evaluations may include compensating errors between different reactions cross-sections (capture, inelastic and/or elastic scattering, and fission), without forgetting the average number of prompt neutrons and the prompt fission neutron spectra. These compensating effects are present in all the evaluations, and the generally-good performance of evaluated data in benchmarks is, to some extent, due to these compensating effects. Nowadays, evaluation efforts are focused on identifying and correcting these effects to restore good performance of the library in integral benchmarks [1]. The development of methodologies and computational tools, beside the benchmarking and validation of new evaluated data will be essential to obtain better accuracies, and to support the design and operation of new nuclear facilities.

The JEFF project has assessed the needs for nuclear data improvements and brings together experts in different areas such as experiments, data evaluation, verification and compilation of the data, processing and benchmarking [2]. In the past, JEFF participants have used different benchmarking suites applied to criticality, fuel inventory, reactivity variation, shielding and activation, decay heat, ... and hence, benchmarking consisted of summing isolated validation results, which resulted in poor coverage of application space. In order to centralize, streamline and strengthen the benchmarking process, the JEFF Benchmarking and Validation Working Group (JEFF-B&V WG) has been created [3].

The NEA Data Bank has taken an increased role in JEFF Project. In addition to carrying out the Project secretariat tasks, and as part of broader nuclear data services to participating countries, the NEA Data Bank also provides a number of services with the JEFF files such as consistency checks, conversion to various pointwise or multigroup formats, file testing and benchmarking using open databases such as ICSBEP and SINBAD. To provide these services the Nuclear Data Evaluated Cycle (NDEC) [4] platform is developed and implemented at the Data Bank.

In this paper, benchmarking efforts from different participating institutions using the general-purpose test libraries JEFF-3.3T1 and T2 are presented. It can be

© The Authors, published by EDP Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
concluded that good performance of the library is achieved.

2. NEA data bank: Tools and databases

NEA Data Bank has supported some recent activities to enhance Nuclear Data Services for participating countries. A comprehensive Q&A process involving verification, testing and benchmarking relies on these NEA tools and databases. These tools and databases have been used to assess the production and selection of JEFF-3.3 file candidates. Hereafter, a brief summary of these tools is presented (see Fig. 1):

- The JANIS (Java-based Nuclear Information Software) software developed by the NEA Data Bank to facilitate the visualization and manipulation of nuclear data, giving access to evaluated nuclear data libraries, such as ENDF, JEFF, JENDL, TENDL etc., and also to experimental nuclear data (EXFOR) and their bibliographical references (CINDA). The on-line JANIS Books provide compilations of cross-sections of evaluated and experimental data from a number of evaluated libraries, nuclear reactions and associated reaction products.

- The Database for the International Criticality Safety Benchmark Evaluation Project (DICE) contains 567 evaluations representing 4874 critical, near-critical, or subcritical configurations into a standardised format that allows criticality safety analysis. This database is easily used to validate calculation tools and perform benchmarking to assess the performance of evaluated nuclear data libraries. DICE also provides the user with access to sensitivity coefficients (percent changes of k-effective due to elementary change of basic nuclear data) for the major nuclides and nuclear processes in a 30-group and 238-group energy structure. These data are currently only available for about 75% of experimental configurations.

- The Nuclear Data Sensitivity Tool (NDaST) is a Java based software, designed to perform calculations on nuclear data sensitivity files for benchmark cases. These calculations are either an estimation of the impact of nuclear data perturbations to the computed case results, and/or; calculation of the uncertainty in the computed results due to evaluated nuclear covariance data. This allows simple and fast analysis for nuclear data evaluators to test the impact of revisions across a wide set of benchmarks. This tool has been recently applied [5] to estimate the impact of improved CIELO files (individual cross-sections, e.g., elastic, inelastic, fission, capture or nu_bar) against the computed or reference case, ENDF/B-VII.1, in around 1500 ICSBEP Benchmarks taken from this Reference [6].

- The Nuclear Data Evaluation Cycle (NDEC) [4] is a systematised workflow for handling and diagnosing the quality of nuclear data under the different steps involved in the production of nuclear data libraries. These steps are the verification, processing, experimental differential validation and experimental integral benchmarking of evaluated nuclear data files. A criticality validation suite of 123 ICSBEP benchmarks is used to assess the reactivity impact of changes to associated nuclear data libraries [7].

3. JEFF benchmarking and validation working group

A challenge for the JEFF Project is to issue a fully consistent library, “complete” with all needed data and associated covariance information, which can be reliably used for a large spectrum of applications, and which has a proven performance level equal or better than that of the present JEFF-3.2 library.

In the JEFF-3.2 evaluation, many benchmarking activities were carried out on criticality, fuel inventory, reactivity variation, shielding and activation, decay heat, . . . However, these benchmarking activities involved summing individual contributions resulting in poor coverage of this evaluation. In addition, the lack of a large, consolidated, reference document for this evaluation resulted in difficulties to reference and an overall decrease in the quality of documentation of this work.

Consequently, JEFF benchmarking activities have been revised to be more coherent and efficient, with the NEA Data Bank taking an increased role in servicing the Project, in particular file testing and benchmarking activities. To fulfill this purpose, the JEFF Benchmarking and Validation (JEFF-B&V) Working Group has been created.

The main goals of JEFF-B&V WG are to exchange ideas on best practices and procedures for benchmarking, and to perform inter-comparisons among different methodologies and validation suites used by JEFF-B&V WG’s participating members. The working group will exchange and cross-check benchmark input decks, discuss benchmarking suites to be used, selection of pertinent validation cases for all nuclides in the library, and provide as complete as possible a multi-purpose validation suite. Institutions such as CIEMAT, ENEA, SCK-CEN, KIT, UKAEA, IRSN, KAERI, JSI, CEA, PSI, JRC, UPM, NRG and NEA have agreed to join efforts to providing thorough benchmarking for upcoming JEFF candidate files.

The centralization of these tasks relies on the NEA Data Bank to implement and co-ordinate a comprehensive process involving verification, testing and benchmarking tasks according to well-defined criteria, while assessing the
needs for benchmarking efforts in participating institutions and helping streamline and rationalize these activities, in particular testing and benchmarking.

Here, the NDEC application [4], under development by NEA, performs the automated testing and diagnosis in human-readable outputs, aiding the selection of better JEFF files. At the same time, DICE/IDAT benchmark databases are used through NdAST to provide a means to easily select benchmarks due to sensitivity coefficients and quickly assessing the impact on benchmark calculations due to nuclear data library changes.

JEFF Project has encouraged users and parties interested in JEFF evaluation to identify the benchmarking data that the new releases should comply with and to provide with the sensitivity analysis to reflect on the performance and deficiencies of the new evaluation.

4. Benchmarking and validation of JEFF-3.3

In this section, the status of benchmarking activities and preliminary results for upcoming JEFF candidate cross-section files are reviewed. The major changes in JEFF-3.3T1/2, with a total number of 559 isotopes, are summarized in Ref. [8]. It has to be remarked the completely new evaluations of 22 isotopes: 235U, 238U, 239Pu, Hf isotopes, etc. it contains unchanged 92 isotopes kept from JEFF-3.2 and 304 new isotopes from TENDL-2015. Feedbacks, bugs and deficiencies in JEFF-3.3 beta version were reported and reviewed by JEFF community.

4.1. General validation

New activities have been initiated at the NEA to perform the validation of JEFF-3.3 against standard, evaluated, microscopic and integral cross-sections. A more thorough analysis has been presented in Ref. [9] as an essential part of activities to validate TENDL library.

4.2. Criticality benchmarks

4.2.1. IRSN – Benchmarking [10]

A first suite of benchmarking activities is created to test and assess new nuclear data evaluations (e.g., 235U, 239Pu and 16O). The DICE database associated with the ICSBEP Handbook is used for this purpose using different criteria for the selection: high sensitivity to new evaluations, low sensitivity to the “background” cross-sections and discharging experiments showing too high experimental uncertainty.

Figure 2 shows the IRSN suite to test new 235U evaluation. Large but discrepant values were observed in HMI-006 series (Fig. 3), part of this effect can be justified due to the change in the 235U capture cross-section in the range 1–100 keV. HM6 series have EALF values in a range between 4–80 keV. Good results are obtained with JEFF-3.3T1/2. Additional analysis in Ref. [11] replacing graphite and copper with other evaluations concluded that these files might have important sensitivity in these cases. Therefore, additional analysis are needed for selecting the best files.

4.2.2. KAERI – Benchmarking [11], NEA – Benchmarking [7]

The second set of validations suites was created to provide a general indication of the overall performance of a given library. In addition, they can help to identify areas where improvements are needed or unexpected discrepancies from changes to nuclear data. The benchmarks are divided according to the fissile material that produces the majority of fissions: Pu, HEU, IEU, LEU, U233 and MIX. Benchmarks are classified by neutron spectra (fast, intermediate and thermal), enrichment, reflector thickness or solution content.

KAERI’s validation calculations were carried out for 119 criticality benchmarks taken from ‘Expanded Criticality Validation Suite’ for MCNP5/6 code packages. JEFF-3.3T1/2 showed good overall performance comparable to ENDF/B-VII.1. Figure 4 shows a better performance of JEFF-3.3.T1/2 than ENDF/B-VII.1 for Pu cases and worse for IEU cases.

NEA’s criticality validation suite contains the 119 KAERI’s cases and 4 additional cases for testing purposes of: 237Np (SMF-8), heavy-water solutions (HST-4), very thermal Pu solution (PST-9 series) and unmoderated ZEUS benchmark (HMF-73) for Cu cross-sections in the fast energy range. Figure 5 shows a good performance of JEFF-3.3.T1 and T2. Calculations are performed with MCNP6.1.
A third type of validation suites is defined to assess specific applications. CEA’s suite is focused on thermal and fast systems (Fig. 6, only for uranium cases). The JEFF-3.3T1 evaluation is compared with JEFF-3.2 on a set of criticality experiments (ICSBEP, CEA/EOLE LWR mockup and CEA/MASURCA SFR mockup, ZPPR) using TRIPOLI-4 code. A general better performance for Pu cases is found. PSI’s suite is focused on LCT (125 cases) and MCT (24 cases) subsets (Fig. 7). Calculations were performed with MCNPX code. Results showed good consistency in LCT cases with JEFF-3.2, and higher deviation in MCT cases due to differences in $^{239}$Pu evaluation.

4.2.4. PSI – Benchmarking [14], NRG [15], IAEA [16] and LANL [17]– Benchmarking

Different expanded validation suites were collected to give a broader overview of the quality of nuclear data libraries. The main contributors to this analysis are PSI (339 cases), CEA (217 cases), NRG (1966 cases), IAEA (429 cases) and LANL(1104 cases, only ENDF/B-VII.1).

Table 1 shows Chi-squared values calculated by IAEA, NRG, LANL and PSI, as each institution uses a different set of cases potential contradictory indications could be obtained. Then, IAEA shows good results for JEFF-3.3T1/2 in Pu and $^{233}$U cases. NRG shows a very high improvement in IEU cases comparing with JEFF-3.1.1. With this information, it can be concluded that JEFF-3.3T1/T2 gives a generally good performance in all systems. For HEU, LEU and IEU improvements in the final JEFF-3.3 release are expected with new upcoming evaluations of $^{238}$U, $^{16}$O and thermal scattering laws for H$_2$O and D$_2$O.

4.3. Shielding and Benchmarks for fusion application

SINBAD database is used as a high quality reference set of Benchmarks for validation of nuclear data used
Table 1. Evaluation of general performance for extended validation suites. Values are Chi-squared and number of cases is in brackets.

| NEA   | NRG   | LANL  | CEA   | PSI   |
|-------|-------|-------|-------|-------|
| E7.1  | 3.3T1 | 3.3T2 | 3.11  | E7.1  | 3.3T1 | 3.3T1 |
| PU    | 4.9 (96) | 3.2 (96) | 4.9 (467) | 4.9 (462) | 6.1 (200) | 2.1 (184) | 1.7 (31) |
| HEU   | 23.1 (160) | 39.9 (160) | 13.4 (727) | 12.7 (698) | 27.7 (425) | 18.2 (6) | 1.6 (70) |
| IEU   | 5.5 (37) | 6.7 (37) | 2.3 (42) | 12.4 (43) | 3.5 (20) | 2.2 (3) | 5.0 (10) |
| LEU   | 1.0 (96) | 1.4 (96) | 4.5 (449) | 3.8 (472) | 2.0 (267) | 5.5 (24) | 3.5 (160) |
| U3    | 2.3 (19) | 1.8 (19) | 4.7 (100) | 3.7 (66) | 4.0 (146) | – | 4.4 (8) |
| MIX   | 2.4 (21) | 3.8 (21) | 7.0 (181) | 5.2 (174) | 2.8 (46) | – | 2.3 (60) |

Figure 9. Comparison of 239Pu cross-sections, fission and capture, between JEFF-3.3T2 and JEFF-3.1.2.

for radiation transport and shielding [18]. New iron cross-section evaluations (56Fe-CIELO and partial new evaluations of 54,57,58Fe) have been used in SINBAD benchmarks recently re-evaluated such as ASPIS IRON-88. The validation revealed general good performance of the new iron data [19,20]. JEFF/Fusion group is working on the reanalysis of the HCPB and HCLL mockup experiments with recent nuclear data evaluations. In general, the current state-of-the-art nuclear data libraries FENDL-3.1b, ENDF/B-VII.1 and JEFF-3.2 showed similar results [21].

Cu-Benchmark experiment irradiated with 14MeV neutrons at the Frascati Neutron Generator at (FNG) is used to check the status of the nuclear data library for copper, in particular the latest release (JEFF-3.3T2), but also earlier data sets such as JEFF-3.1.1/3.2 and FENDL-3. A comparison of C/E values of irradiated activation foils at different penetration depths, and neutron and gamma-ray spectra were performed. The main conclusions of this work are: 1) underestimation up to 15% is found for high threshold reactions using JEFF-3.1.1 and FENDL-3 and up to 20% using JEFF-3.2, good improvement in JEFF-3.3T2; 2) underestimation of the low threshold reaction is more severe (30%), furthermore, a decreasing trend versus penetration depth is observed, JEFF-3.2 provides the largest underestimation; 3) non-threshold reactions show even larger underestimation (60%), with the C/E decreasing as a function of the penetration depth, all the libraries produce the same results. It was concluded that results of the Cu experiment call for a deep revision/re-evaluation of the copper cross-sections [22,23].

Within the JEFF Fusion WG two extensions to the MCNP code have been developed, MCS4EN to carry out sensitivity analysis, and MCUNED to handle deuteron data libraries in the transport simulation. Extensive usage, validation and testing of these tools on well-defined benchmarks have been performed [24,25].

4.4. Burnup benchmarks

The performance of nuclear reaction data should be assessed in different depletion scenarios using powerful burnup/depletion codes [26]. A PWR burnup pin-cell benchmark proposed by UAM [27] has been used to calculate differences between evaluations. The keff produced with JEFF-3.3T1 underestimates JEFF-3.1.2’s results by ~500 pcm (at 60 GWd/t) [28]. This effect might be mainly due to the new evaluation of 239Pu. More 239Pu disappears along burnup with JEFF-3.3T1 compared to JEFF-3.1.2; in parallel, the products of 239Pu capture (240Pu, 241Pu and 242Pu) appear in larger concentrations with JEFF-3.3T1. The one-group integrated 239Pu capture cross-section with JEFF3.3T1 is ~1.5% larger than the value with JEFF3.1.2.

Finally, fuel assembly decay heat benchmarks are a valuable piece of information to test long lived products and actinides. Preliminary results in Ringhals 2&3 measurements for UO2 assemblies are presented in Ref. [29], JEFF-3.3.1T calculated values over predict by 3–4% with respect to JEFF-3.2.

4.5. Activation benchmarks

A review of current decay data evaluations has shown the importance to identify and fill pressing gaps. Particularly for decay heat calculations where beta decay and gamma contributions play an important role (e.g., TAGS, spectra or missing components of the spectra – particularly at high energy gamma, . . .) [30].

4.6. Other nuclear data validation

Several activities focused on validation of effective kinetics parameters ($\beta$eff and $\lambda$) based on IRPhE experiments [31] are still in progress, more effort is needed to analyse the effect of switching from 6 to 8 families of neutron precursors.

4.7. Reactor benchmarks and new designs

New reactor concept designs such as ASTRID (Advanced Sodium Technological Reactor for Industrial Demonstration) and MYRRHA (Multi-purpose hYbrid Research Reactor for High-tech Applications) have been used for testing purposes. In ASTRID, the impact of new 235Pu and 239Pu files is assessed. Pb and Bi files are tested in MYRRHA design [32].

4.8. Covariance verification

NDaST tool [33] can be used to carry out a comparison of nuclear data uncertainty propagated from covariance files. An example of it can be found in Ref. [34].
Sensitivity and uncertainty analysis for the Tritium Breeding Ratio (TBR) of a DEMO Fusion Reactor was carried out with different sets of covariance data: JEFF-3.2, “JEFF-3.2+TENDL-2014” and JEFF-3.3T1. The resulting relative STDs of the total TBR uncertainties are 3.2%, 8.6% and 5.6%, respectively, dominated by the 16O due to the high sensitivity to 6Li-TBR [35, 36].

In the framework of the JEFF project, new methodologies to propagate nuclear data covariance have been developed. SANDY code [37] developed by SCK-CEN has demonstrated high potential to assess the response uncertainty in different nuclear systems.

Table 2. Few examples of calculated uncertainties (in pcm) in [34] from covariance files.

| Case Label | JEFF-3.3T1 | ENDF/B-VII.1 | JENDL-4.0 | Exp. Uncert. |
|------------|------------|--------------|-----------|-------------|
| PCI031-001 | 547        | 656          | 1100      |             |
| HMF030-001 | 630        | 882          | 100       |             |
| IMF033-001 | 1278       | 729          | 170       |             |
| HST032-001 | 231        | 230          | 260       |             |
| HM006-001  | 1996       | 644          | 80        |             |
| PM001-001  | 780        | 542          | 129       |             |

where covariance data of JEFF-3.3T1, ENDF/B-VII.1 and JENDL-4.0 are tested in the NEA criticality validation suite. As it was expected, large differences in keff uncertainties were found due to 239Pu and 235U covariance data in JEFF-3.3T1 (see Table 2).

It can be concluded that the first assessment of JEFF-3.3T1/2 nuclear data library has shown a good performance. Although, more efforts for improving files such as Cu are still needed. New benchmarking activities are foreseen to assess the impact of new evaluations such as 16O-IRSN and 238U-CIELO files.

This paper collects the Benchmarking and Validation activities of many institutions involved in JEFF project and presented within JEFF/B&V-Working Group. This paper acknowledges those joint efforts providing thorough studies and analysis to improve JEFF files.

References

[1] M. Chadwick, NDS 118, 1–25 (2014)
[2] JEFF Co-ordination Group, JEF/DOC-1730 (2015)
[3] O. Cabellos, JEF/DOC-1729 (2015)
[4] F. Michel-Sendis, JEF/DOC-1625 (2015)
[5] I. Hill, SG40-15, Ian HILL (2016)
[6] S. Marck, NDS 113(12), 2935–3005 (2012)
[7] O. Cabellos, JEF/DOC-1756 (2016)
[8] D. Rochman, JEF/DOC-1732 (2015)
[9] M. Fleming, JEF/DOC-1723 (2015)
[10] R. Ichou, JEF/DOC-1753 (2016)
[11] D.H. Kim, JEF/DOC-1755 (2016)
[12] P. Tamagno, JEF/DOC-1752 (2016)
[13] M. Pecchia, JEF/DOC-1760 (2016)
[14] D. Rochman, JEF/DOC-1754 (2016)
[15] S. van der Marck, private communication
[16] A. Trkov, private communication
[17] S. Kahler, private communication
[18] I. Kodeli, JEF/DOC-1725 (2015)
[19] A. Trkov, EFFDOC-1272 (2015)
[20] I. Kodeli, EFFDOC-1295 (2016)
[21] A. Klix, EFFDOC-1294 (2016)
[22] M. Angelone, EFFDOC-1276 (2015)
[23] A. Klix, EFFDOC-1280 (2015)
[24] K. Kondo, EFFDOC-1230 (2014)
[25] I. Kodeli, EFFDOC-1276 (2015)
[26] A. Stankovskiy, JEF/DOC-1728 (2015)
[27] K. Ivanov et al., NEA/NSE/DOC (2012) I
[28] L. Fiorito, JEF/DOC-1758 (2016)
[29] W. Haeck, JEF/DOC-1750 (2016)
[30] M. Fleming, EFFDOC-1293 (2016)
[31] P. Leconte, JEF/DOC-1722 (2015)
[32] F. Álvarez, JEF/DOC-1756 (2016)
[33] J. Dyrda, JEF/DOC-1727 (2015)
[34] J. Dyrda, JEF/DOC-1759 (2016)
[35] E. Nunnenmann, EFFDOC-1275 (2015)
[36] E. Nunnenmann, EFFDOC-1290 (2016)
[37] L. Fiorito, JEF/DOC-1726 (2015)
[38] T.T Ivanova et al. Nucl. Sci. Eng. 145(1), 97–104 (2003)