Expanded COG criticality validation suite for inter-laboratory benchmark data comparison

David Heinrichs[1*], *Soon* Kim[1], and *Scott* McKinley[1]

[1]Lawrence Livermore National Laboratory, Livermore, California, U.S.A.

Abstract. The COG suite of criticality benchmarks has been formally expanded from 591 to 3,395 to cover the entire energy range from thermal to fast neutron spectra under a variety of reflector and moderator conditions and fissile materials. COG results have been compared with benchmark values from the International Criticality Safety Benchmark Evaluation Project Handbook for ENDF/B-VII.1, ENDF/B-VIII.0 and JEFF-3.3. COG results have been also compared with a MERCURY validation suite. Most of the results agreed with the benchmark values within ±3σ. Among the three cross section data, cases with ENDF/B-VIII.0 performed best with about 85% of the total cases within ±3σ range. A major inter-comparison project between COG, MCNP, MORET, and SCALE for ENDF/B-VIII.0 and JEFF-3.3 is in progress.

1 Introduction

COG [1] is a general purpose, multi-particle, high-fidelity Monte Carlo code developed by LLNL. Since 2017, LLNL has focused on expanding COG benchmark cases as part of a collaborative effort of the benchmark inter-laboratory comparison study between the US Depart of Energy (DOE) Nuclear Criticality Safety Program (NCSP) and the French Institut de Radioprotection et de Sûreté Nucléaire (IRSN). The benchmark cases fully cover the entire range from thermal to fast neutron spectra for a wide variety of fissionable material forms in a variety of reflector and moderator conditions described in the International Criticality Safety Benchmark Evaluation Project (ICSBEP) Handbook [2].

The original number of LLNL 591 benchmark cases (143 PU, 358 U-235, and 90 U-233) was expanded to 3,395 PU, HEU (Highly Enriched Uranium), IEU (Intermediate Enriched Uranium), LEU (Low Enriched Uranium), U-233, Mixed fuel, and SMF (Special Metal Fast) cases, providing valuable data for the inter-laboratory benchmark data comparison**.** The number of benchmark cases in each of these six major categories is summarized in Table 1.

The cross section data libraries used are ENDF/B-VII.1, ENDF/B-VIII.0, and JEFF-3.3. Calculations for ENDF/B-VIII.0 were performed with: (a) continuous-energy cross sections based on ENDF/B-VIII.0 nuclear data as processed by the International Atomic Energy Agency (IAEA); (b) probability tables for the unresolved resonance region as processed by Brookhaven National Laboratory using NJOY within the ADVANCE system [3]; and (c) thermal scattering laws using algorithms developed by LLNL [4]. The most recent version, COG11.3, was used for all the benchmark cases.

Table 1. Number of benchmark cases.

Category	Number of Cases
PU	766
HEU	1,056
IEU	207
LEU	807
U-233	193
Mixed	356
SMF	10
Total	3,395

2 Results

2.1 Comparison with ICSBEP benchmarks

Tables 2, 3 and 4 summarize results for the 3,395 benchmark cases. Most of the cases agree with the benchmark values and uncertainties within ±3σ. 668 cases with ENDF/B-VII.1, 492 cases with ENDF/B-VIII.0 data and 573 cases with JEFF-3.3 data exceed this range. For PU, ENDF/B-VIII.0 and JEFF-3.3 performed better than ENDF/B-VII.1. For HEU, ENDF/B-VIII.0 performed better than ENDF/B-VII.1 and JEFF-3.3. COG results with ENDF/B-VIII.0 for LEU and IEU were better than those with JEFF-3.3 and ENDF/B-VII.1 data. For U-233, ENDF/B-VIII.0 performed better than ENDF/B-VII.1 and JEFF-3.3. For Mixed cases, JEFF-3.3 performed better than ENDF/B-VII.1 and ENDF/B-VIII.0.

---

[*] Corresponding author: heinrichs1@llnl.gov

Table 2. COG results compared with benchmark values for ENDF/B-VII.1.

SD	< 1σ	1σ-2σ	2σ-3σ	> 3σ	Total
PU	295	174	147	150	766
HEU	544	165	116	231	1,056
IEU	137	28	11	31	207
LEU	345	171	114	177	807
U-233	119	39	16	19	193
MIXED	195	61	40	60	356
SMF	7	1	2	0	10

Table 3. COG results compared with benchmark values for ENDF/B-VIII.0.

SD	< 1σ	1σ-2σ	2σ-3σ	> 3σ	Total
PU	484	137	53	92	766
HEU	576	181	100	199	1,056
IEU	157	38	3	9	207
LEU	391	183	117	116	807
U-233	123	37	18	15	193
MIXED	174	85	36	61	356
SMF	5	3	2	0	10

Table 4. COG results compared with benchmark values for JEFF-3-3.

SD	< 1σ	1σ-2σ	2σ-3σ	> 3σ	Total
PU	488	139	47	92	766
HEU	541	180	108	227	1,056
IEU	145	28	12	22	207
LEU	383	165	89	170	807
U-233	111	46	16	20	193
MIXED	184	83	47	42	356
SMF	3	2	3	2	10

To compare performance of each set with the benchmark data, the root mean square errors (RMSE) are also calculated. The RMSE is defined as,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(K_{c,i}-K_{b,i})^2}{N}}$$

where N is the total number of cases, $K_{\text{c,i}}$ and $K_{\text{b,i}}$ are the calculated and benchmark k_{eff} values, respectively. This represents a sample standard deviation of the differences between calculated and the benchmark values.

The RMSE for the six categories are compared in Fig. 1. ENDF/B-VIII.0 data performed better than others for PU, HEU, LEU, and Mixed categories.

The χ^2 (chi-squared) value is the indicator in determining the degree of difference between the calculated and the benchmark values. Here, χ^2 is defined as:

$$\chi^2=\frac{1}{\nu}\sum_{i=1}^{N}\frac{(K_{c,i}-K_{b,i})^2}{\sigma_i^2}$$

where $K_{\text{c,i}}$ and $K_{\text{b,i}}$ are the calculated and benchmark k_{eff} values, respectively. ν is the degree of freedom (DoF), and σ_i^2 is the variance. Fig. 2 shows the cumulative χ^2 values of the three different cross section data for six categories of the expanded COG criticality validation suite.

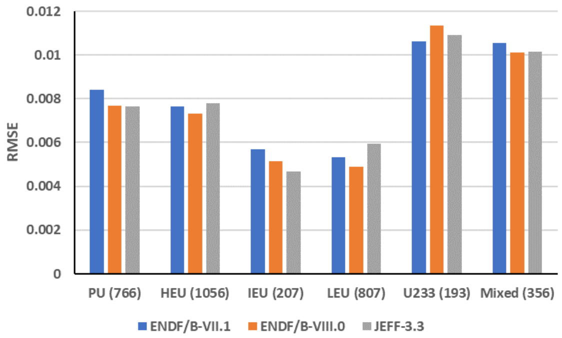

Fig. 1. RMSE for ENDF/B-VII.1, ENDF/B-VIII.0, and JEFF-3.3.

In Fig. 2, '71' in the parentheses represents ENDF/B-VII.1, '80' for ENDF/B-VIII.0, and 'J33' for JEFF-3.3, respectively. Cumulative χ^2 values for the six categories are summarized in the figure comparing performance of ENDF/B-VII.1 (dashed line), ENDF/B-VIII.0 (solid line), and JEFF-3.3 (dotted line). Steep slopes in the curves are the cases where COG overpredict or underpredict the benchmark values significantly. The initial steep slope in the IEU (71) curve (dotted black lines) is due to the overprediction of IEU-MET-FAST-020 cases (20% ^{235}U Metal). The similar steep slope in the PU curves (red lines) is largely due to the PU-COMP-MIX (Plutonium-oxide polystyrene cubes) cases. The benchmark uncertainties in these cases may be under-estimated, or there may be additional sources of experimental uncertainty not considered in the evaluation. The U-233 cases also show a steep slope, which raises questions as to the quality of many of the solution experiments that date back to the 1950s.

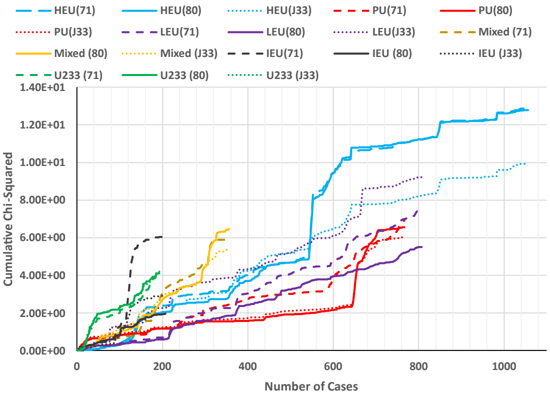

Fig. 2. Cumulative χ^2 for Six Benchmark Categories.

In general, COG results with ENDF/B-VIII.0 show better agreement with benchmark values than with ENDF/B-VII.1 or JEFF-3.3. Possible sources of discrepant results come from: (a) ENDF/B-VII.1, ENDF/B-VIII.0 or JEFF-3.3 nuclear data; (b) additional errors associated with processing the nuclear data; (c) errors in modeling the benchmarks; and (d) errors in the experimental benchmark measurements themselves and their evaluated biases and uncertainties.

2.2 Comparison with MERCURY validation suite

To validate the newly expanded benchmark data set, COG input decks are translated into the MERCURY [5] input decks. From this effort, errors from the benchmark models, if any, cannot be identified. However, results from different (MERCURY) cross section data processing can be compared. The selected 3,350 COG benchmark cases matching the corresponding MERCURY validation suite based on ENDF/B-VIII.0 are compared.

For Mercury, calculations using ENDF/B-VIII.0 data were performed with: (a) continuous energy cross sections based on ENDF/B-VIII.0 nuclear data as processed by LLNL's FUDGE into the GNDS format; (b) thermal neutron scatter law data as processed with FUDGE; and (c) probability tables for the unresolved resonance region were not been included. The most recent version, Mercury 5.25.0-137 along with GIDI 3.19.65, was used for all benchmark cases. The COG to Mercury translation was done with software still in testing associated with Mercury 5.25 series. Table 5 summarizes categorized benchmark comparison cases.

Table 5. Number of MERCURY Benchmark Cases.

Category	Number of Cases
PU	753
HEU	1,034
IEU	207
LEU	797
U-233	193
Mixed	356
SMF	10
Total	3,350

COG and MERCURY results with varying standard deviations against the benchmark data are compared in Tables 6. Out of 3,350 cases, results of about 85% agree with the benchmark values and uncertainties within ±3σ. 492 cases with COG, 507 cases with MERCURY exceeded this range. For PU and HEU cases, COG performed slightly better than MERCURY. The reverse is true for Mixed cases.

Table 6. COG and MERCURY result comparison.

SD	$< 1\sigma$		1σ - 2σ		2σ - 3σ		$> 3\sigma$	
Code	COG	MERCURY	COG	MERCURY	COG	MERCURY	COG	MERCURY
PU	475	391	134	184	52	68	92	110
HEU	570	560	180	177	98	107	186	190
IEU	157	140	38	40	3	5	9	22
LEU	390	425	178	148	113	107	116	117
U-233	123	120	37	40	18	13	15	20
MIXED	174	190	85	87	36	31	61	48
SMF	5	4	3	4	2	2	0	0

3 Conclusions

COG11.3 comparison with benchmark values from the ICSBEP Handbook and MERCURY results showed quite good agreement with each other. Sources of discrepant results may come from 1) errors in the cross section data, 2) possible errors from the modeling of the benchmark experiments, or 3) errors in the benchmark measurement data itself or its evaluated biases and uncertainties.

An inter-laboratory comparison project with different Monte Carlo codes such as MCNP, MORET, and SCALE for ENDF/B-VIII.0 and JEFF-3.3 is in progress. LLNL participation in this project will result in development of significantly more COG benchmark cases as our goal is to overlap the VALID, WHISPER, and IRSN compendia of criticality benchmarks to the extent possible, which will be beneficial to international code user communities.

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and was funded by the U.S. Department of Energy Nuclear Criticality Safety Program.

4 References

1. *COG: A Multiparticle Monte Carlo Code Transport Code, Version 11.1*, CCC-829, Released by RSICC (2015).
2. *International Handbook of Evaluated Criticality Safety Benchmark Experiments*, NEA/NSC/DOC (95)03, OECD, NEA, Paris, France (December 2019 Edition).
3. D. Brown, "Use of a Continuous Integration and Deployment Software to Automate Nuclear Data V&V," BNL (2014).
4. D. E. Cullen et al., "Thermal Scattering Law Data: Implementation and Testing using the Monte Carlo Neutron Transport Codes COG, MCNP and TART" (May 2003).
5. P. Brantley, M. McKinley, "Mercury Web Site," (2021). [Online]. Available: https://wci.llnl.gov/simulation/computer-codes/mercury/.