LA-UR-17-20453

Title: LANL-SNL Collaboration on NCS Validation

Author(s): Brown, Forrest B.
Miller, John
Henderson, Shawn
Rising, Michael Evan
Alwin, Jennifer Louise

Intended for: ANS 2017 Annual Meeting, 2017-06-11/2017-06-15 (San Francisco, California, United States)

Issued: 2017-01-23

LANL-SNL Collaboration on NCS Validation

Forrest Brown[1], John Miller[2], Shawn Henderson[2], Michael Rising[1], Jennifer Alwin[1]

[1] *Los Alamos National Laboratory, PO Box 1663, Los Alamos, NM, 87544, fbrown@lanl.gov*
[2] *Sandia National Laboratory, PO Box 5800, Albuquerque, NM, 87185, millerj@sandia.gov*

INTRODUCTION

During 2016, nuclear criticality safety (NCS) practitioners from SNL and code developers from LANL collaborated in several areas of interest to the DOE/NNSA Nuclear Criticality Safety Program (NCSP). This collaboration involved

- Testing of the preliminary release of the MCNP6-Whisper methodology, with feedback to the developers,
- Sharing of the benchmark catalogs (the collection of MCNP input files and benchmark results), with 1101 cases from LANL and 866 cases from SNL,
- Comparison and analysis of 357 benchmarks common to both catalogs,
- Investigation of the impact of the different benchmark catalogs on sensitivity-uncertainty based NCS validation results from MCNP6-Whisper,
- Investigation of the impact of randomized selections from the benchmark catalog on sensitivity-uncertainty based validation results from MCNP6-Whisper.
- Investigation of the use of MCNP6-Whisper in selecting benchmarks for use in NCS validation for unique, nonstandard, legacy fuel applications.

This paper summarizes the collaboration work and initial results. It must be noted that the results described herein are preliminary and need further research and detailed analysis. However, the initial results are very interesting, and it is important to share them with the NCSP community.

DESCRIPTION OF COLLABORATION EFFORT

MCNP6-Whisper Methodology

Whisper [1-7] is a statistical analysis package that supports NCS validation. It uses the sensitivity profile data for an application as computed by MCNP6 [8] along with covariance files [6,7] for the nuclear data. The application sensitivity profile is matched against a precomputed catalog of benchmark sensitivity profiles, selecting the most neutronically-similar benchmarks, and then statistical analysis determines the bias and bias uncertainty. Portions of the margin for subcriticality (MOS) due to nuclear data and code are estimated, and a baseline upper-subcritical-limit (USL) is determined. NCS analysts may then use this baseline USL from Whisper along with their expert judgment to help set the USL for a range of normal and credible abnormal conditions. Whisper-1.0 [1-3] was developed in 2014 and used exclusively at LANL. Whisper-1.1 [4-7] is the portable version developed during 2015-2016 for distribution with the release of MCNP6.2.

Part of the LANL-SNL collaboration was “friendly testing” of the MCNP6-Whisper package by an independent organization (SNL). Feedback from the first external users at SNL was important for both software quality assurance (SQA) and code usability. All results described below were obtained using MCNP6.1.1 [9], Whisper-1.1, and ENDF/B-VII.1 [10] cross-section data.

Sharing of Benchmark Catalogs

In traditional NCS validation, a set of benchmark experiments from the ICSBEP handbook [11] is selected to cover a range of expected NCS applications. In new sensitivity-uncertainty based NCS validation, a very large set of benchmarks is used, along with the Whisper techniques for automatically selecting the benchmarks from the catalog that are most similar to an application.

The SNL benchmark catalog consisted of 866 cases from ICSBEP, with 265 of those from [12] and 601 others that are currently being reviewed. The LANL benchmark catalog consisted of 1101 cases from ICSBEP, with 1086 from [3] and 15 new additions [5].

As part of the sharing, all of the SNL benchmarks were updated to use ENDF/B-VII.1 nuclear data, material modeling based on isotopes (rather than elemental data), and the recommended best practices for running modern Monte Carlo codes [13] (i.e., 10,000 or more neutrons/cycle, use Shannon entropy checks to ensure that the fission distribution has converged).

Comparison of Benchmark Catalogs

Comparing the benchmark catalogs showed that 357 cases were common to both. These 357 cases were used in analyses described below, to assess the quality of the benchmark catalogs and to investigate potential “analyst bias.” Different analysts at different sites set up the 357 common benchmarks independently based on ICSBEP specifications. The benchmarks were run using the same code, the same nuclear data, and the same Monte Carlo control parameters (e.g., neutrons/cycle, discarded cycles, active cycles). Any bias determined from the 357 common cases would suggest differences due to analyst modeling procedures.

Impact of Benchmark Catalogs on USL Results

The MCNP6-Whisper methodology was carried out 4 different ways using the SNL & LANL benchmark catalogs: (1) using only the LANL catalog, (2) using only the SNL catalog, (3) using the LANL catalog with additions from the SNL catalog (excluding common cases), (4) using the SNL catalog with additions from the LANL catalog (excluding common cases). Any differences in the computed baseline USLs would serve to quantify the impact of different benchmark catalogs on NCS validation and the impact of combining the catalogs into one much larger catalog.

Impact of Randomized Benchmark Catalogs

Further initial work was performed to investigate the impact of the size and completeness of the benchmark catalog on the baseline USL results determined by MCNP6-Whisper. For several application problems, Whisper was run multiple times, with 50% of the benchmark catalog randomly excluded from each run. Differences in the baseline USLs from the repeated cases give an indication of how sensitive the Whisper methodology is to the precise contents of the benchmark catalog.

RESULTS FROM COLLABORATION WORK

Comparison of Benchmark Catalogs

If the bias and bias uncertainty are computed separately for the 357 cases common to both LANL and SNL suites, using ordinary statistics and the normality assumption, the overall agreement is excellent. Table 1 shows the results along with the average difference in SNL and LANL cases. The average difference in results is less than 1 standard deviation.

Table 1. Comparison of common benchmarks

	Bias ± Std.Dev
357 LANL cases	0.00175 ± 0.00024
357 SNL cases	0.00179 ± 0.00021
Ave. SNL-LANL	0.00004 ± 0.00010

Differences between Keff results for the 357 cases are summarized in Table 2 in terms of absolute differences:

The 18 cases with $|\Delta k| > 0.001$ are being examined to determine modeling differences or errors. In some cases, there are modeling differences such as including or not including impurities, or using a simplified model vs. a detailed model. Detailed comparisons are in progress.

Table 2. Distribution of differences in common cases

range for $|k_{SNL}-k_{LANL}|$	Number of cases
< 0.001	339
0.001 – 0.002	8
0.002 – 0.003	3
0.003 – 0.004	1
0.004 – 0.005	1
> 0.005	5

Impact of Benchmark Catalogs on USL Results

Individual and combined SNL and LANL benchmark catalogs were used with Whisper to perform and compare preliminary USL calculations for a single case from current NCS work at SNL. Benchmark catalogs were constructed and exclusion files were built using Whisper's native statistical rejection method.

The benchmark catalogs used include SNL's and LANL's individual catalogs and combinations of both. The total combined catalog includes 1610 cases. The 1610 cases were built in two different manners. One (LANL+SNL) keeps LANL's cases that are common between sets and the other (SNL+LANL) keeps SNL's. Table 3 shows USLs for the application case calculated by Whisper for each catalog.

Table 3. Catalog USL calculations

Benchmark Catalog	Highest C_k	Bias+Bias Uncertainty	Calculated USL
SNL	0.9902	0.01624	0.97747
LANL	0.9924	0.01715	0.97656
SNL+LANL	0.9924	0.01691	0.97680
LANL+SNL	0.9924	0.01691	0.97680

These results agree very well with each other, although SNL's individual catalog resulted in a slightly higher USL. The other three catalogs include benchmarks that characterize the selected case more closely. These benchmarks have a greater combined bias and bias uncertainty that resulted in a very small difference and lower USL.

Impact of Randomized Benchmark Catalogs

Seven applications problems related to the LANL PF4 Facility were chosen. Each of these was run with Whisper 25 times using the LANL benchmark catalog. For each of the 25 runs for a case, 50% of the benchmark cases were selected randomly and excluded from the Whisper calculations. The minimum, average, and maximum of the 25 USLs for each case are shown in Table 4.

Table 4. Whisper results for 25 repetitions using random 50% of benchmark catalog

Application	Whisper baseline USL		
	Min	Ave	Max
1 - Pu cylinder, H/D=0.5	0.978	0.979	0.981
2 - Pu cylinder, H/D=3.0	0.977	0.979	0.980
3 - Pu annulus, H/D=1.0, IR=0.001	0.978	0.979	0.981
4 - Pu annulus, H/D=1.0, IR=0.5	0.978	0.979	0.981
5 - Pu – NaCl sphere, 1" water refl	0.975	0.977	0.978
6 - Pu sphere, 0.01 cm Ta reflector	0.978	0.979	0.981
7 - Pu sphere, 5.0 cm Ta reflector	0.924	0.929	0.933

- All cases used 4500 g of Pu-239
- Cases 1-4 used 2.54 cm radial water reflector
- Cases 3-4 used water in center of annulus
- Cases 6-7 used spherical Pu with Ta reflector

For cases 1-6, the effect on baseline USL from randomly excluding 50% of the benchmark catalog is negligible. For those 6 cases, Whisper was able to find sufficient numbers of similar benchmarks with high correlation coefficients, and the statistical analysis results for USLs were very close. For case 7, however, none of the benchmarks in the catalog had correlation coefficients in the 0.8-0.9 range, and the Whisper analysis indicates that the benchmark catalog is not complete enough. In fact, none of the benchmarks in the LANL catalog included thick tantalum reflectors, so this result is expected. Even so, the range in predicted baseline USLs for case 7 is significant, but not large.

Benchmark Selection for Nonstandard Applications

In attempting to validate the SNL calculational methods for analyzing uranium-gadolinium fuel for some legacy applications, Whisper was used to assess the similarity of the SNL benchmark catalog to the applications. It was discovered that the traditional SNL validation suite of the time did not adequately cover the neutronics of the applications, that the correlation coefficients from Whisper were low. To remedy this, 77 additional benchmarks containing gadolinium were found in ICSBEP and added to the SNL benchmark catalog. This effort will be reported in detail separately.

SUMMARY & FUTURE WORK

The benefits from the collaboration described in this summary include:

- Additional SQA, testing, and checking of the respective benchmark catalogs.
- Identification of specific benchmarks that warrant further detailed review.
- The combined effort eases the task of expanding the benchmark catalogs for use in NCS validation.
- Feedback from independent, external testing of a new software package (Whisper) strengthens the usability and SQA. Lessons-learned can be dealt with prior to the official public release of the software.
- Initial results of the comparisons suggest that no apparent "analyst bias" is present between the NCS validation work at the respective sites. That is, even when different analysts at different sites independently construct the benchmark models, the resultant baseline USLs determined by MCNP6-Whisper show only small effects.
- Different sets of benchmarks in the benchmark catalogs have only very small effects on the baseline USLs determined by the MCNP6-Whisper methodology. Even randomly excluding 50% of the benchmarks in the catalog has little effect on resultant baseline USLs. The MCNP6-Whisper method for selecting relevant, neutronically-similar benchmarks for the statistical analysis is robust and not very sensitive to the precise contents of the benchmark catalog (assuming that it is large).

The LANL-SNL collaboration work to date has benefitted both sites, and both are interested in continuing this work. As noted in the introduction, much of this work is preliminary, in the initial stages. The preliminary results to date suggest a number of worthy areas for additional collaboration:

- Expand both benchmark catalogs to include more cases and provide more complete coverage of possible applications.
- Perform more real-world application testing on the use of MCNP6-Whisper based NCS validation, including comparisons with traditional NCS validation methods.
- Perform further detailed analysis using the different benchmark catalogs, to thoroughly investigate the notion of "analyst bias." If such bias exists, it would dictate additional margin needed in the MOS for validation. The preliminary conclusion, however, is that no evidence of "analyst bias" is apparent.
- Review and comment on work and results to date, addressing conformance with ANS-8.24 and other standards.
- Explore the use of the MCNP6-Whisper methodology for applications where there are not a sufficient number of ICSBEP benchmarks available.

ACKNOWLEDGMENTS

The LANL work was supported and encouraged by the US DOE/NNSA Nuclear Criticality Safety Program. We are indebted to NCSP for their insistence on getting things right.

REFERENCES

1. B.C. Kiedrowski, F.B. Brown, et al., “Whisper: Sensitivity/Uncertainty-Based Computational Methods and Software for Determining Upper Subcritical Limits”, *Nucl. Sci. Eng.* [also LA-UR-14-26558] (2015).
2. B.C. Kiedrowski, "User Manual for Whisper (v1.0.0), Software for Sensitivity- and Uncertainty-Based Nuclear Criticality Safety Validation", LA-UR-14-26436 (2014).
3. B.C. Kiedrowski, et al., "Validation of MCNP6.1 for Criticality Safety of Pu-Metal, -Solution, and -Oxide Systems", LA-UR-14-23352 (2014).
4. F.B. Brown, M.E. Rising, J.L. Alwin, “User Manual for Whisper-1.1”, LA-UR-17-xxxxx (2017).
5. F.B. Brown, M.E. Rising, J.L. Alwin, “Release Notes for Whisper-1.1”, LA-UR-17-xxxxx (2017).
6. R.C. Little, T.Kawano, et al., “Low-fidelity Covariance Project”, *Nuclear Data Sheets* **109**, 2828-2833 [also LA-UR-08-7732] (2008).
7. F.B. Brown, M.E. Rising, “Covariance Data File Formats for Whisper-1.0 & Whisper-1.1”, LA-UR-17-20098 (2017).
8. J.T. Goorley, et al., "Initial MCNP6 Release Overview," *Nuclear Technology*, **180**, 298 (2012).
9. J.T. Goorley, "MCNP6.1.1-Beta Release Notes," LA-UR-14-24680 (2014).
10. M.B. Chadwick, et al., “ENDF/B-VII.1 Nuclear Data for Science and Technology: Cross Sections, Covariances, Fission Product Yields and Decay Data,” LA-UR 11-05121, Los Alamos National Laboratory, also: *Nuclear Data Sheets*, **112** Issue 12, pg. 2887-2996 (Dec 2011).
11. International Handbook of Evaluated Criticality Safety Benchmark Experiments, NEA No. 7231, OECD Nuclear Energy Agency (2014).
12. K.R. DePriest, J.A. Miller, “Creation of a Benchmark Suite for SNL Nuclear Criticality Safety Code Validation”, SAND2015-9476, Sandia National Laboratory (2015).
13. F.B. Brown, “A Review of Best Practices for Monte Carlo Criticality Calculations”, ANS NCSD-2009, Richland, WA, Sept 13-17 paper LA-UR-09-3136, presentation LA-UR-09-5623 (2009).