ICNC 2019 - 11th International conference on Nuclear Criticality Safety
September 15-20, 2019 – Paris, France

TOOLS FOR VALIDATION AND UNCERTAINTY QUANTIFICATION WITH ANSWERS SOFTWARE

Paul Smith [(1)*], David Hanlon [(1)], Geoff Dobson [(1)], Magda Stefanowska [(1)], Simon Richards [(1)], Richard Hiles [(1)], Christophe Murphy [(1)]

[(1)] ANSWERS Software Service, Wood, Kings Point House, Queen Mother Square, Poundbury, Dorchester, Dorset, DT1 3BW, United Kingdom.
** paul.smith6@woodplc.com*

ABSTRACT

The MONK® categorisation scheme has long been available to MONK users to assist in the choice of appropriate validation experiments for a chosen application. This has recently been supplemented by the production of a similarity index which indicates how similar an experiment is to the chosen application in terms of the sensitivity of the multiplication factor to the nuclear data. These tools have been used to select appropriate experiments for the OECD NEA wetted MOX powder benchmark. The results indicate that it is essential to use expert judgement in addition to the tools to ensure that appropriate experiments are chosen.

This paper also describes the production of a prototype Bayesian updating validation tool for use with ANSWERS® software. This is also applied to the OECD NEA benchmark on wetted MOX powders and the results for the first six cases are described. It is shown how the validation tool can be used to refine the estimate of the neutron multiplication factor and its uncertainty. For the six application cases described it is shown that up to a factor of two reduction in the estimated uncertainty can be obtained by the use of validation data.

KEY WORDS

ANSWERS, validation, uncertainty.

1. INTRODUCTION

In the ICNC 2015 conference Santamarina reported that the OECD Working Party on Nuclear Criticality Safety (WPNCS) Expert Group on Experimental Needs had identified a paucity of experimental data representative of low-moderated MOX powders [1]. He proposed an NEA benchmark exercise based around six wetted MOX powder application cases [1]. An essential part of the exercise is the process by which appropriate validation experiments are chosen for the application cases. The exercise was expanded to include an additional nine application cases by Carmouze *et al*. [2].

The application cases defined in the benchmark exercise each consist of a sphere of MOX surrounded by a spherical shell of water. In each case 3 wt% H_2O is present and the radius of the MOX sphere, Pu concentration and Pu vector are as displayed in Table 1, for the first six cases to be studied here.

The aim of the exercise is to estimate the multiplication factor, k_{eff}, and uncertainty on this value for the application cases. Validation experiments are then chosen to allow the user to estimate the bias in their chosen code system for the application cases and thus revise the estimate of the multiplication factor and its associated uncertainty.

The chosen code system for the calculations reported here is that developed by Wood Plc's ANSWERS® Software Service. The ANSWERS Software Service [3] has developed a suite of software for criticality, reactor physics, shielding and dosimetry applications. ANSWERS codes are widely used in over thirty countries around the world, and on a range of reactor types including: AGR, BWR, CANDU, MAGNOX,

RBMK, PBMR, PWR, VVER and many experimental reactors. The primary tool for criticality safety analysis within the ANSWERS suite of software is the MONK® Monte Carlo criticality and reactor physics code.

Table I. Benchmark Specification for Cases 1-6 [1]

Case	MOX Radius (cm)	M_{PuO2}/M_{MOX} (%)	^{239}Pu/Pu (%)	^{240}Pu/Pu (%)	^{241}Pu/Pu (%)	^{242}Pu/Pu (%)
1	17.0	100	71	17	11	1
2	22.5	30				
3	46.0	12.5				
4	17.7	100	64	23	10	3
5	24.1	30				
6	52.5	12.5				

2. MONK Overview

a. MONK

MONK is a 3D Monte Carlo code for criticality analysis and reactor physics calculations [4, 5]. MONK's advanced geometry modelling allows exact representation of complex geometries. The FG (fractal geometry) geometry option provides combinatorial geometry using basic solid bodies (cuboids, ellipsoids, cylinders, cones, etc.) and can be combined with HG (hole geometry) representations of more complicated shapes using Woodcock tracking [6]. Geometry can also be imported from CAD files using several different techniques including tracking through NURBS, to provide a faithful representation of the CAD geometry. The FG, HG and CAD geometries can be combined flexibly so that different parts of the geometry can be represented by the different geometric constructs.

MONK's superhistory algorithm [7] greatly reduces the correlation between the different stages of the Monte Carlo calculation thus providing robust and reliable estimates of the neutron multiplication factor and its associated uncertainty. Superhistory powering also improves coverage of the domain in each stage, thus improving convergence for loosely coupled systems.

b. Nuclear Data Libraries

MONK's BINGO collision processor and associated BINGO continuous energy nuclear data libraries provide accurate simulation of neutronic behaviour. BINGO nuclear data libraries exist for MONK from a variety of sources including JEFF, ENDF/B, and CENDL. The BINGO libraries contain data for around 300 nuclides, allowing a wide range of nuclear systems to be modelled.

BINGO libraries contain continuous energy nuclear data with cross-sections represented as a smoothly varying function of energy and tabulated on an energy grid that allows linear interpolation with a maximum interstitial error in the cross-section of 0.1%. The evaluated data are processed using the BINGO pre-processor (BPP) code, in order to transform the evaluated data from the ENDF-6 format to a form that allows efficient Monte Carlo sampling. Point cross-sections are generated using an external processing 'engine', usually the NJOY99.364 RECONR and BROADR modules. The outgoing angle and energy data for neutrons are transformed into equi-probable bins or probability functions. The S(α, β) thermal scattering data are transformed to ease bivariate sampling in momentum and energy, allowing MONK to generate secondary energy data for emitted neutrons during the run. In the resolved resonance range, the point energy representation is adequate to model resonance shielding effects. In the unresolved resonance range, a small number of nuclides, including ^{232}Th, ^{233}U, ^{238}U and ^{239}Pu, require processing with the sub-group method. Here, group data are generated at infinite dilution and at an appropriate level of resonance self-shielding. Pairs of sub-groups are formed and adjusted to preserve the infinite dilute and shielded cross-sections. This

effectively moves the resonance peaks to one of the sub-groups forcing a flux depression dependent on the level of shielding, avoiding the need to sample cross-sections in the unresolved resonance region.

A covariance data library has been generated for use with MONK. The covariance library can be used in the quantification of nuclear data uncertainties by application in conjunction with the sensitivity methods available in the code as well as in the generation of sampled nuclear data libraries. The NJOY modules ERRORJ and COVR were used to process the evaluated covariance data into BOXER format in 44 energy groups. The group scheme is a revised version of the SCALE 6.1 44-group scheme, with some group boundaries adjusted to be coincident with the standard WIMS 172-group bounds.

The covariance library contains what was judged the best available data for each nuclide (as of September 2015). The data were taken from JEFF-3.2, ENDF/BVII. 1, JENDL-4.0 and TENDL-2011. A small number of nuclides used data from JEFF-3.1.2, JENDL-3.3 or CENDL-3.1. Data taken from JEFF-3.2 include ^{1}H, ^{10}B, C and ^{56}Fe. Data taken from ENDF/B-VII.1 include Zr, ^{235}U, ^{238}U and ^{239}Pu.

c. MONK Validation Database

Validation is of prime importance for criticality and reactor physics codes and the MONK package contains a large set of validation data, drawing heavily on experiments in the International Criticality Safety Benchmark Evaluation Project (ICSBEP) Handbook [8] and the International Reactor Physics Experiment Evaluation (IRPhE) Project Handbook [9].

The validation database for MONK comprises critical experiments covering a wide range of conditions regarded as being representative of systems that are widely encountered in the nuclear industry. From each experiment, a selection of the most important configurations has been modelled for inclusion in the database. At the time of the MONK10B release in 2017 there were 801 experimental configurations in the database. For each of these, MONK10B calculations were carried out using eight different nuclear data libraries (BINGO format: JEF2.2, JEFF3.1.2, JEFF3.2, ENDF/B-VII.1 and CENDL3.1; and DICE format: JEF2.2, ENDF/B-VI.3 and JENDL3.2). For each combination of experimental configuration and data library, results from four independent MONK calculations were combined to achieve a standard error on k_{eff} of 0.0005. All of the results are provided to users in the MONK User Guide [10].

In addition to a summary of the benchmarks and results published in the MONK User Guide, each experiment in the MONK validation database has an accompanying validation report describing the detailed specification of the benchmark, an analysis of the results and comparisons with other codes. A rigorous, independent peer review process is followed before a new benchmark is accepted into the MONK validation database, giving users considerable confidence in the quality of the validation evidence available to them. Therefore, while it might be relatively simple to model additional benchmarks in MONK, it requires significant time and resources to add them formally to the validation database.

For this reason, a new *Tier 2* validation database has been created, allowing the available validation evidence for MONK to be expanded more rapidly. The Tier 2 validation cases have not yet been through the full peer review process required of the Tier 1 validation set, but have been self-checked to a level consistent with the example calculations in the ICSBEP Handbook [8] and can provide very useful additional validation evidence. Those Tier 2 validation benchmarks which are most applicable to current user validation requirements will be targeted for elevation to Tier 1 status.

To date 1145 Tier 2 validation cases have been added, all of which were taken from the ICSBEP Handbook [8], more than doubling the number of available validation benchmarks and taking the total to 1946. The calculation/experiment results for the Tier 1 and Tier 2 validation benchmarks are shown in Figure 1.

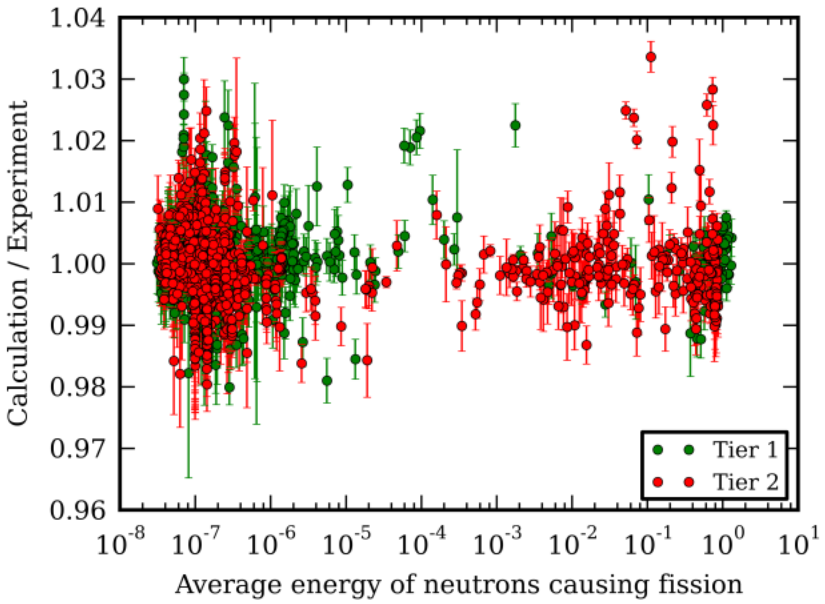

Figure 1. Calculation/Experiment results for Tier 1 and Tier 2 benchmarks plotted against the logarithmic average energy of neutrons causing fission. Error bars show the combined experimental and calculation uncertainties.

3. Prior Estimates

The prior uncertainties were estimated by using the sensitivity option in MONK combined with the covariance data stored on the BINGO nuclear data library. The covariance data is stored on the library in the form:

$$C_{ij} = \frac{COV(X_i,\ X_j)}{X_i X_j} \tag{1}$$

where X_i is the i^{th} item of nuclear data, e.g. cross-section in a particular energy group for a particular reaction in a particular nuclide. The sensitivity coefficients are defined as:

$$S_i = \frac{X_i}{k}\frac{\partial k}{\partial X_i} \tag{2}$$

The prior estimate of uncertainty for application case α is then given by:

$$\frac{\sigma_{\alpha,\text{prior}}}{k} = \sqrt{S_{\alpha i} C_{ij} S_{\alpha j} + \sigma^2_{\alpha,\text{MC}}} \tag{3}$$

where $S_{\alpha i}$ is the sensitivity coefficient of application case α to nuclear data item i, and summation over indices i and j is implied. The Monte Carlo simulation variance, $\sigma^2_{\alpha,\text{MC}}$, can be made small enough to be neglected by running sufficient samples.

4. Choosing Appropriate Validation Experiments

Two tools are available with MONK to help the user in the choice of appropriate validation experiments: a Similarity Index and the MONK categorisation scheme. These tools provide guidance for the user, however, the analyst must apply intelligent reasoning in the final selection of validation experiments, i.e. expert judgement is required.

a. Similarity Index

A similarity tool has been developed that evaluates two sensitivity indices. The one used here, E_{sum}, is defined by [11]:

$$E_{\text{sum}} = \frac{\boldsymbol{B} \cdot \boldsymbol{S}}{|\boldsymbol{B}||\boldsymbol{S}|} \tag{4}$$

where $\boldsymbol{B}$ is the vector of sensitivities in the application case and $\boldsymbol{S}$ is the vector of sensitivities in the chosen experiment. This index is used rather than the correlation coefficient as it is independent of the nuclear data covariance matrix which is subject to change when the evaluated libraries are updated. The similarity indices for the application cases and a small selection of experiments are shown in Table II. Only experiments having a similarity index greater than 0.7 for at least one of the application cases are shown. Note that setting the threshold for acceptance at 0.7 results in between four and six experiments being chosen for each of the application cases.

Table II. Similarity Indices > 0.7

Experiment\Application Case	1	2	3	4	5	6
PU-MET-FAST-037-013	0.86			0.84		
PU-MET-FAST-037-013-002	0.76			0.75		
PU-MET-FAST-002-001	0.73			0.72		
PU-MET-FAST-001-001-014	0.72			0.71		
PU-COMP-MIXED-002-006		0.76	0.74		0.76	0.75
MIX-COMP-INTER-005-001		0.74	0.90		0.76	0.93
MIX-MISC-MIXED-001-007		0.83	0.85		0.84	0.86
MIX-MET-INTER-003-001					0.70	
PU-COMP-MIXED-001-002		0.75	0.73		0.76	0.74
PU-MET-FAST-001-001	0.73			0.71		
MIX-MISC-MIXED-001-002		0.71			0.71	

b. MONK Categorisation Scheme

Six system parameters are defined to categorise the neutronic behaviour of the system [10]:

A. Type of fissile material (3 partitions: uranium, plutonium, other).
B. Non-fuel absorption (3 partitions: low, medium, high).
C. Leakage (3 partitions: low, medium, high).
D. Resonance absorption (3 partitions: low, medium, high).
E. Fast fission (3 partitions: low, medium, high).
F. Hydrogen fuel content (4 partitions: none, low, medium, high).

In this way a six digit integer, the category number, is assigned to every system. When a MONK calculation is performed the category number is automatically calculated and displayed in the code output. The category numbers for all of the experiments in the MONK validation set are listed in the MONK User Guide [10].

Experiments chosen using the categorisation scheme are displayed in Table III. Note that three of the experiments identified using this scheme have similarity indices less than 0.7 and so were not selected using the Similarity Index. The categorisation scheme was modified to produce a real number for each system parameter rather than an integer value 1, 2 or 3. This overcomes the problem associated with values that are close to the boundaries but on opposite sides.

Table III. Experiments suggested by the Categorisation Scheme & their Similarity Indices

Case	Experiment	Similarity
1	PU-MET-FAST-037-013	0.86
2	PU-COMP-MIXED-002-006	0.76
3	MIX-COMP-INTER-005-001	0.9
	MIX-MET-FAST-008-001	< 0.70
	MIX-MET-INTER-004-001	< 0.70
	MIX-MISC-MIXED-001-007	0.87
4	None	-
5	MIX-COMP-INTER-005-001	0.76
	MIX-MET-INTER-004-001	< 0.70
6	MIX-MISC-MIXED-001-002	< 0.70
	MIX-MISC-MIXED-001-007	0.86
	MIX-COMP-INTER-005-001	0.93
	MIX-MET-FAST-008-001	< 0.70
	MIX-MET-INTER-004-001	< 0.70

From Table III it is seen that there are no experiments identified as appropriate for Case 4. Also it will be seen below that the experiment PU-COMP-MIX-002-006 identified as appropriate for Case 2 is deemed unsuitable on grounds of expert judgement, so additional data must also be sought for this case. In order to identify additional potentially appropriate experiments the criteria must be relaxed. So Table IV displays experiments that differ from the application cases in only one of the six system parameters.

From Table IV it is seen that three experiments are identified for Case 2, one of which (MIX-MISC-MIXED-001-007) was suggested by the similarity index, with a value of 0.83. Two experiments are also identified for Case 4, one of which (PU-MET-FAST-037-013) was suggested by the similarity index, with a value of 0.84. Thus combining the experiments in Tables III and IV provides cover for all six application cases.

Table IV. Experiments suggested by the Categorisation Scheme – 1 different system parameter

Case	Experiment	Similarity
1	None	-
2	MIX-MET-FAST-008-001	< 0.70
	MIX-MET-INTER-004-001	<0 .70
	MIX-MISC-MIXED-001-007	0.83
3	PU-COMP-MIXED-002-006	0.74
4	PU-MET-FAST-037-013	0.84
	PU-COMP-MIXED-002-006	< 0.70
5	PU-COMP-MIXED-002-006	0.76
	MIX-MET-FAST-008-001	< 0.70
	MIX-MISC-MIXED-001-007	0.84
	MIX-MET-INTER-003-001	0.70
6	PU-COMP-MIXED-002-006	0.75

c. Expert Judgement

The MONK categorisation scheme and the similarity index provide guidance on the choice of suitable experiments for validation, however, it is essential that the user applies experience and intelligent reasoning in the final selection of experiments. In this instance an experienced criticality expert was consulted for advice on the experiments selected using the similarity index and the MONK categorisation scheme.

Regarding experiments PU-COMP-MIXED-001-002 and PU-COMP-MIXED-002-006, similar experimental configurations have been reported and they all over-predict k_{eff} with all nuclear data libraries; the reason for this is not known. A number of modifications were made to the system definition by the experimenters in order to simplify the modelling. It is not known whether these original adjustments applied by the experimenters are the reason that the calculations do not produce the expected value of k_{eff}. This raises concerns over the inclusion of these experiments in the validation set for the specified applications. Therefore the analysis has been performed twice; once with these experiments included and then again with them excluded. The two affected experiments are highlighted in grey in Tables II, III and IV and are not included in the analysis presented in Section 5 except for the initial calculations in order to determine the effect of their inclusion. It can be seen from Table 5 that the calculations performed for this study for experiments PU-COMP-MIXED-001-002 and PU-COMP-MIXED-002-006 also over-predicted the reported k_{eff} values as observed in previous studies. This resulted in a high estimated bias that is not supported by comparison with the other experiments.

Note that if a BEPU (best estimate plus uncertainties) approach is used the inclusion of the above two experiments would result in the subtraction of a large bias that may not be justified and may result in a significant under-prediction of k_{eff}. However, if a conservative approach is adopted in which negative bias adjustments are not made, than the use of the above experiments would have no effect on the estimated value of k_{eff}.

5. Posterior Estimates

For updating the prior estimates the multiplication factor and its uncertainty we use the Bayesian updating scheme employed in AREVA's MOCABA code [12]. The estimated bias for Case α, $k_{\alpha,\text{bias}}$ is given by [13]:

$$k_{\alpha,\text{bias}} = S_{\alpha i} C_{ij} S_{\varepsilon j} \left[S_{\varepsilon k} C_{kl} S_{\delta l} + \tilde{C}_{\varepsilon\delta} \right]^{-1} \frac{\Delta k_\delta}{k_\delta} \tag{5}$$

where C_{ij} /C_{kl} is the covariance matrix for the nuclear data and $\tilde{C}_{\varepsilon\delta}$ is the covariance between experiments ε and δ resulting from manufacturing tolerances (i.e. uncertainties in dimensions, compositions and detectors). For this analysis $\tilde{C}_{\varepsilon\delta}$ is taken to be a diagonal matrix with the diagonal elements being the variances of the experiments as reported in the ICSBEP documentation. For a discussion of how to proceed if the experimental covariance matrix is not diagonal see [14].

The posterior uncertainty is given by:

$$\sigma^2_{\alpha,\text{post}} = \sigma^2_{\alpha,prior} w - S_{\alpha i} C_{ij} \tilde{S}_{\varepsilon j} \left[\tilde{S}_{\varepsilon k} C_{kl} S_{\delta l} + \tilde{C}_{\varepsilon\delta} \right]^{-1} S_{\alpha m} C_{mn} \tilde{S}_{\delta n} \tag{6}$$

where $\tilde{S}_{\varepsilon j}$ is the sensitivity coefficient of experiment ε to nuclear data item j, and $S_{\alpha i}$ is the sensitivity coefficient of application case α to nuclear data item i (as in Section 3).

Note that this implies that the addition of validation data always results in a reduction in the estimated uncertainty. However, if the uncertainty in the experimental data (the term in square brackets in Equation 6) is large, the inverse of this term will be small and hence the reduction will be small. Similarly the term that appears in front of and behind the term in square brackets represents the degree of correlation between the

application case and each of the experiments. If this is small, the reduction in the estimated uncertainty will also be small.

6. Results

Initial estimates of bias and posterior uncertainty were performed using all of the experiments as shown in Table II. The results are displayed in Table V. This shows that there are large estimated biases in Cases 2, 3, 5 and 6 and that the estimated uncertainty is several hundred pcm in all cases. The final two columns show the equivalent results when experiments PU-COMP-MIXED-002-006 and PU-COMP-MIXED-001-002 are excluded from the analysis. This affects the results for cases 2, 3, 5 and 6 where it is seen that the exclusion of the above experiments significantly reduces the estimated bias in the results; the effect on the estimated uncertainty is small.

Table V. Posterior estimates: based on Similarity Index with experiments PU-COMP-MIXED-002-006 and PU-COMP-MIXED-001-002 Included and Excluded

Case	Prior Uncertainty (pcm)	Bias - Included (pcm)	Posterior Uncertainty - Included (pcm)	Bias - Excluded (pcm)	Posterior Uncertainty - Excluded (pcm)
1	526	-154	432	-154	432
2	561	-1222	427	-596	460
3	519	-974	300	-692	310
4	516	-157	427	-157	427
5	517	-1119	385	-605	409
6	564	-858	282	-695	287

From this point all calculations were performed with experiments PU-COMP-MIXED-002-006 and PU-COMP-MIXED-001-002 excluded from the analysis.

Table VI provides a comparison of results obtained from experiments chosen using the Similarity Index with those chosen based on the MONK categorisation scheme. It is seen that the results compare well and show consistency between the two schemes.

Table VI. Posterior estimates: Comparison of Similarity Index (SI) and MONK Categorisation Scheme (MCS)

Case	Prior Uncertainty (pcm)	Bias - SI (pcm)	Posterior Uncertainty – SI (pcm)	Bias – MCS (pcm)	Posterior Uncertainty - MCS (pcm)
1	526	-154	432	-61	482
2	561	-596	460	-633	403
3	519	-692	310	-680	297
4	516	-157	427	-74	438
5	517	-605	409	-742	354
6	564	-695	287	-573	278

Finally Table VII shows a comparison of results using the experiments selected using the Similarity Index, as displayed in Table II, plus those suggested by the MONK categorisation scheme, displayed in Table III. This doubles the number of experiments used for Case 3 from two to four and for Case 6 from three to six. This results in a small reduction in the estimated bias and uncertainty in both cases. For case 5 one experiment is

added to the list of four previously used, resulting in a small reduction in the estimated bias and a negligible reduction in uncertainty.

Overall it is seen that the use of validation data allows biases in the calculated results to be reduced and reduces the uncertainty in the estimated multiplication factors. In Case 6 the estimated uncertainty is reduced by a factor of two.

Table VII. Posterior estimates: Comparison of Similarity Index and Both SI & MCS

Case	Prior Uncertainty (pcm)	Bias - SI (pcm)	Posterior Uncertainty – SI (pcm)	Bias - MCS (pcm)	Posterior Uncertainty – MCS (pcm)
1	526	-154	432	-154	432
2	561	-596	460	-596	460
3	519	-692	310	-680	297
4	516	-157	427	-157	427
5	517	-605	409	-615	409
6	564	-695	287	-584	274

7. CONCLUSIONS

ANSWERS has developed a number of tools to help the code user to estimate the uncertainty on calculated results. These tools are being implemented in Visual Workshop, an integrated development environment (IDE) for ANSWERS Software. The tools include a similarity index to aid in the choice of suitable experiments for validation. The MONK categorisation scheme is another aid to the choice of experiments and has been available to MONK users for many years. A prototype Bayesian updating tool for validation (based on MOCABA [12]) has been developed to allow users to use experimental data to refine estimates of the neutron multiplication factor and its associated uncertainty. This results in a reduction in the estimated uncertainty when new experimental data is added. The reduction is small if there is large uncertainty on the measured data or the experiment is poorly correlated with the application case.

The tools have been applied to the OECD NEA benchmark on wetted MOX powder [2, 3]. This has emphasised the need for expert judgement in addition to the similarity index and MONK categorisation scheme when choosing appropriate validation experiments. The results show how the validation tool can be used to refine the estimate of k_{eff} and how the use of validation data can reduce the estimated uncertainty, by up to a factor of two for the application cases chosen.

REFERENCES

[1] A. Santamarina and C Riffard, “The MOX Powder Calculation Improvement: OECD Benchmark Proposal and CRISTAL-RIB Results,” *Proceedings of ICNC 2015*, Charlotte USA, 14-17 September, pp.1043-1052 (2015).

[2] C. Carmouze, J. Huyghe and A. Santamarina, “MOx Powders Calculation Improvement: Criticality Calculations in the context of OECD NEA benchmark,” *Proceedings of GLOBAL 2017*, Seoul Korea, 24-29 September, (2015).

[3] http://www.answerssoftwareservice.com

[4] S. D. Richards, C. M. J. Baker, P. Cowan, N. Davies and G. P. Dobson, “MONK and MCBEND: Current Status and Recent Developments”, *Annals of Nuclear Energy*, **82**, pp. 63-73 (2015).

[5] S. Richards, G. Dobson, T. Fry, D. Hanlon, D. Long, R. Perry, P. Smith, F. Tantillo and T. Ware, "Recent Developments to the MONK Monte Carlo Code for Criticality Safety and Reactor Physics Analyses", *Proceedings of ICNC 2019, Paris, France* (2019).

[6] E. Woodcock, T. Murphy, P. Hemmings and S. Longworth, "Techniques used in the GEM Code for Monte Carlo Neutronics Calculations", *Proceedings of Conference on the Applications of Computing Methods to Reactor Problems*, ANL-7050, pp. 557-579 (1965).

[7] R. J. Brissenden and A. R. Garlick, "Biases in the Estimation of *k*eff and its Error by Monte Carlo Methods", *Annals of Nuclear Energy*, **13**, pp. 63-83 (1986).

[8] "International Handbook of Evaluated Criticality Safety Benchmark Experiments", NEA-1486/16, NEA Nuclear Science Committee (2018).

[9] "International Handbook of Evaluated Reactor Physics Benchmark Experiments", NEA-1765/13, NEA Nuclear Science Committee (2018).

[10] "MONK – A Monte Carlo Program for Nuclear Criticality Safety and Reactor Physics Analyses. User Guide for Version 10B", ANSWERS/MONK/REPORT/013, Answers Software Service (2017).

[11] B. L. Broadhead *et al*., "Sensitivity and Uncertainty-Based Criticality Safety Variation", *Nuclear Science and Engineering*, **146**, pp. 340-366 (2004).

[12] A. Hoefer, O. Buss, M. Hennebach, M. Schmid and D. Porsch, "MOCABA: A General Monte Carlo - Bayes Procedure for Improved Predictions of Integral Functions of Nuclear Data", *PHYSOR 2014, Kyoto Japan* (2014).

[13] B. L. Broadhouse, B. T. Rearden, C. M. Hopper, J. J. Wagschal, C. V. Parks, "Sensitivity- and Uncertainty-Based Criticality Safety Validation Techniques", *Nuclear Science and Engineering, 146, 340-366 (2004)*

[14] Maik Stuke, Axel Hoefer, Oliver Buss, Maksym Chernykh, Geoff Dobson, James Dyrda, Tatiana Ivanova, Nicolas Leclaire, William J Marshall, Dennis Mennerdahl, Bradley Rearden, Paul Smith, Fabian Sommer, Sven Tittelbach, "UACSA Phase IV: Role of Integral Experiment Covariance Data for Criticality Safety Validation Summary of Selected Results", *ICNC 2019, Paris, France, September 15-20, (2019).*