ASSESSMENT OF PERFORMANCE OF THE BAYESIAN BASED UNCERTAINTY ANALYSIS APPROACH VS. THE FREQUENTIST STATISTICS APPROACH IN APPLICATION TO CRITICALITY SAFETY

A. Vasiliev[1], H. Lee[1]*, M. Frankl[1], M. Hursin[1,2], D. Rochman[1], H. Ferroukhi[1]
[1]Paul Scherrer Institut, Forschungsstrasse 111, 5232 Villigen PSI, Switzerland
[2]École Polytechnique Fédérale de Lausanne, Switzerland

alexander.vasiliev@psi.ch, matthias.frankl@psi.ch, mathieu.hursin@epfl.ch, dimitri-alexandre.rochman@psi.ch, hakim.ferroukhi@psi.ch

ABSTRACT

In the present work a comparison of two distinct approaches on the interpretation of validation results for light water reactor (LWR) fuel criticality assessments is presented: one based on the frequentist tolerance limits and another on the Bayesian framework. In general, both the frequentist statistics and the Bayesian model approaches have their intrinsic advantages and disadvantages and, therefore, it is relevant to compare results of criticality safety evaluations (CSE) obtained with both methods. Of particular interest in this context is the application of CSE with the Burnup Credit (BUC) concept for LWR used nuclear fuel (UNF), the composition of which differs from the fresh fuel compositions primarily available worldwide for validation studies. This paper is oriented towards providing an illustration on the comparison of different CSE variants in application to a model of an UNF disposal canister filled with identical fuel assemblies, as a function of burnup. In the performed study it was found that the Bayesian approach provides less penalizing results, leading in the considered example to some relaxation of the burnup requirement for the UNF criticality safety, namely by ~2.5MWd/kg for the case of the 5wt% initially enriched pressurized water reactor (PWR) fuel. However, the behavior of the Bayesian based results is counterintuitive – the safety margins become less penalizing as burnout increases, although no benchmarks with used nuclear fuel are available in the employed validation suite. Such performance of the utilized methods should be further verified in future enhanced studies.

1. INTRODUCTION

In the field of nuclear criticality safety (NCS) there exist well established procedures for the definition of safety criteria on the basis of appropriate validation studies (i.e., a comparison of calculation results with evaluated experimental measurements). Many methodologies are traditionally based on the use of the frequentist statistics concepts, such as tolerance intervals. There also exists a comprehensive collection of evaluated criticality benchmark experiments, namely the International Criticality Safety Benchmark Evaluation Project (ICSBEP) Handbook, which facilitates required validation studies. The conventional frequentist inference approach assumes that the benchmark experiments selected for validation studies represent well an application system of interest.

*Current affiliation: KAERI, Korea

There also exist methodologies based on the Bayesian inference[3] [1],[2], which can adjust calculation results using the information on similarity between an application case and the analysed benchmarks. Such information may be incorporated in the frequentist-based methodologies too, e.g. as concerns nuclear data uncertainty propagation. Thus, advanced CSE methodologies require knowledge of the correlations between the criticality calculation results (i.e. the effective neutron multiplication factor, k_{eff}) for an application case and the benchmarks, as well as the correlations between the benchmarks themselves. In general, the origins of the correlations can be related to the commonality of the system designs, measurement techniques or, in particular, due to the employed code and nuclear data libraries.

The given work is an extension of the previous studies on the subject of CSE and BUC realised at PSI with modern calculation capabilities, in particular as concerns the propagation of nuclear data uncertainties [3],[4]. On the one hand, recently, a new activity has been proposed within the OECD/NEA/NSC/WPNCS, to assess the level of maturity and reliability of available BUC methodologies for applications related, for instance, to final disposal canisters loaded with UNF [5]. The motivation was related to the fact that existing CSE methodologies, including their validation bases, may have different levels of conservatism and comprehensiveness, depending on the application type. Some methodologies can be based on "generic" Nuclear Criticality Safety (NCS) criteria, while other "case-specific" criteria can be defined, for instance, involving adjusted Nuclear Data (ND) or an adjusted calculation bias and the related uncertainty for an application system k_{eff}, based on either frequentist or Bayesian statistics (or even their combination). The optimal level of complexity of the methodology is still an open question presently, in particular in the context of potential industrial applications such as NCS of UNF final disposal canisters. For example, the use of a single generic Upper Subcritical Limit (USL) would simplify the derivation of the loading curves for UNF noticeably. Case-specific (e.g. Burnup (BU)-dependent USL) approaches, instead, can be more efficient for reducing the number of required disposal canisters (or for savings on the canister design). The effects of such potential methodological differences on the loading curves, in terms of required minimal BU vs. initial fuel assemblies' enrichment, are not obvious. Therefore, OECD/NEA/NSC/WPNCS SG13 exercise was proposed to examine such effects by a direct comparison of the participants' results for a well-defined and simplified pseudo-application case [5],[6].

On the other hand, it was also recently observed at PSI that the use of the covariance matrices (CM) in the calculated to experimental (C/E) validation results' evaluations may lead to implausible results [7]. An assumption was made in [7] that the observed results belong to the 'PPP'-type anomaly [8],[9] although further verification and analysis are required. The main concern, which also served as a motivation for this work, was whether the considered CM might similarly lead to problematic results if applied with a Bayesian updating framework, such as described in [10].

For the reader's convenience, Section 2 of this paper provides a concise summary on the employed CSE methodologies. As pseudo application cases, test models of the Swiss UNF canister corresponding to Nagra's preliminary design [11] were used. Section 3 summarises results of relevant verification studies (here the verification of the calculation procedures and tools is meant), to justify correctness of the applied methodology realisations. Section 4 presents the main results obtained in this work and finally conclusions are provided in Section 5.

[3] Interested readers can find robust definitions for the frequentist and Bayesian inferences, for instance, in [1]. Furthermore, for a brief comparison of the Bayesian philosophy to frequentist philosophy in application to nuclear fuel systems CSE, work [2] can be recommended.

2. BRIEF DESCRIPTION OF THE APPLIED METHODOLOGIES

2.1 Standard CSE Approach for Fresh LWR Fuel

The current reference PSI CSE approach was presented in detail in [12],[7]. The NCS criterion is defined as

$$k_a^{cal} + \Delta k_a^{cal} < USL = k^{LTB} - \Delta k^{AM} - \Delta k^{AOA}\,, \tag{1}$$

where k_a^{cal} and Δk_a^{cal} are an application case k_{eff} value and its calculation uncertainty; Δk^{AM} is an "administrative margin" to guarantee sub-criticality, and Δk^{AOA} is an additional margin if the application system is outside the validation area of applicability (AOA).

The k^{LTB} parameter is a lower tolerance bound (LTB) which can be estimated based on the sample of the calculated-to-benchmark k_{eff} values, $k_i = \frac{k_i^{cal}}{k_i^{bench}}$.

The Upper Subcritical Limit (USL) in such approach is generic and can be applied to any LWR fuel configuration. The case of BUC is not explicitly considered in this work and the required modifications of (1) to be applicable for BUC, according to PSI's current methodology, were reported in [11]. For better illustration of the Δk_a^{cal} uncertainty feature, Eq. (1) can be re-written as

$$k_a^{cal} + \Delta k_a^{cal} - k^{LTB} < -\Delta k^{AM} - \Delta k^{AOA}. \tag{2}$$

In the following, the value of Δk^{AM} is assumed to be 0.05. The component Δk^{AOA} is out of scope of the given work and in the following is set to zero. The k_{normal}^{LTB} value can be determined using the assumption of normality:

$$k_{normal}^{LTB} = \overline{k} - k_1(\alpha, p, N) \times s, \tag{3}$$

or without any distribution assumption, using order statistics [12]. Parameters $\overline{k}$ and s are defined as

$$\overline{k} = \frac{\sum_{i=1}^{N} w_i k_i}{\sum_{i=1}^{N} w_i}\,,\; s = \sqrt{\frac{\frac{1}{(N-1)}\sum_{i=1}^{N} w_i (k_i - \overline{k})^2}{\frac{1}{N}\sum_{i=1}^{N} w_i}}\,, \tag{4}$$

where the weighting factor w_i is defined by $\frac{1}{\sigma_i^2}$, which includes the Monte Carlo statistical uncertainty (σ_i^{MC}, usually negligible), optionally - the ND-related uncertainty (σ_i^{ND}) [3], and the experimental benchmark uncertainty (σ_i^{bench}):

$$\sigma_i = k_i \sqrt{\left(\frac{\sigma_i^{MC}}{k_i^{cal}}\right)^2 + \left(\frac{\sigma_i^{ND}}{k_i^{cal}}\right)^2 + \left(\frac{\sigma_i^{bench}}{k_i^{bench}}\right)^2}. \tag{5}$$

In this case, less reliable benchmarks contribute less in the determination of k_{normal}^{LTB} and vice versa. Note that following the proposal made in [3], the nuclear data uncertainties of the calculated k_{eff} values were also included in the weighting procedure in the present work, though it should be mentioned that, for the given study, their contribution in Eq. (4) did not make strong changes to the final results.

In case of the order statistics, k_{NP}^{LTB} value is defined as

$$k_{NP}^{LTB}(\alpha,p,N) = \min_{l(\alpha,p,N)}\left(k_{i=1,N}\right), \tag{6}$$

$$l(\alpha,p,N) = \underset{l\in[1,N]}{\operatorname{argmax}}\left(1-\sum_{i=0}^{l-1}\frac{N!}{i!(N-i)!}(1-p)^i p^{N-i} \geq 1-\alpha\right). \tag{7}$$

The parameters k_a^{cal} and k^{LTB} are in fact random variables dependent on the ND library (NDL). The uncertainty of the difference $(k^{LTB}-k_a^{cal})$ depends on the correlations r (the Pearson correlation coefficients) between the application system and the k^{LTB} value:

$$\sigma_{\left(k^{LTB}-k_a^{cal}\right)}^{ND} = \sqrt{VAR\left(k^{LTB}-k_a^{cal}\right)} = \sqrt{VAR(k^{LTB})+VAR\left(k_a^{cal}\right)-2COV\left(k^{LTB},k_a^{cal}\right)} =$$

$$= \sqrt{\left(\sigma_{k^{LTB}}^{ND}\right)^2+\left(\sigma_{k_a^{cal}}^{ND}\right)^2-2r_{\left(k^{LTB},k_a^{cal}\right)}\sigma_{k^{LTB}}^{ND}\sigma_{k_a^{cal}}^{ND}}\,. \tag{8}$$

The application case uncertainty Δk_a^{cal}, in general, consists of the ND-unrelated, such as technological/manufactural tolerances, and ND-related components:

$$\Delta k_a^{cal} = n\sqrt{\sigma_{\left(k_a^{cal}-k^{LTB}\right)}^{not\ ND\ 2}+\sigma_{\left(k_a^{cal}-k^{LTB}\right)}^{ND\ 2}}\,. \tag{9}$$

In the case k_{eff} is determined by means of Monte Carlo calculations, the statistical uncertainty should be taken into account in $\sigma_{\left(k_a^{cal}-k^{LTB}\right)}^{not\ ND}$. However, in the modern practice it normally has a negligible value compared, e.g., to the ND-related uncertainties. In this paper only the ND-related uncertainties will be considered and thus Δk_a^{cal}=n$\sigma_{\left(k^{LTB}-k_a^{cal}\right)}^{ND}$. In the following, the case of n=2 is considered. The value of $\sigma_{\left(k^{LTB}-k_a^{cal}\right)}^{ND}$ can be calculated with the PSI in-house tool NUSS [13],[4]. The use of the randomly sampled ND generated with NUSS allows evaluation of the correlation and respectively covariance matrices for parameters of interest, e.g., the application case and benchmarks uncertainties [14].

2.2 Bayesian Model Approach

An alternative to the above presented frequentist approach is the Bayesian based approach, which is well described, e.g., in [15]. Without repeating the details presented in [15], here only the most relevant equations used by the Bayesian model (BM) methodology are outlined. The Bayesian concept assumes adjustment of the application case calculated k_a^{cal} value and its uncertainty based on the validation results and the stronger the correlations between an application system and the benchmark are, the more significant can be the adjustment. Similarly to the previous section, for the sake of simplicity only the ND-related uncertainties for the application case are considered in this work. The adjusted $k_a^{cal^*}$ and its uncertainty then can be obtained with the following equations [15]:

$$k_a^{cal^*} = k_a^{cal} + \Sigma_{ab}U^{-1}\Delta k^{bench}\,; \tag{10}$$

$$\sigma_{\left(k_a^{cal^*}\right)} = \sigma_{\left(k_a^{cal}\right)} - \Sigma_{ab}U^{-1}\Sigma_{ab}^T, \tag{11}$$

where $\Delta k_i^{bench} = \left(k_i^{bench} - k_i^{cal}\right)$ and Σ_{ab} represents the vector of the covariances between an application case and the validation benchmarks. Matrix U contains the covariances between the validation benchmark k_i^{cal} values, which consist of the two parts: one related to the nuclear data covariances and another related to the covariances due to the benchmark specifications:

$$U = U^{ND} + U^{bench};\ \ U_{i,j}^{ND} = \rho_{i,j}^{ND}\sigma_i^{ND}\sigma_j^{ND};\ \ U_{i,j}^{bench} = \rho_{i,j}^{bench}\sigma_i^{bench}\sigma_j^{bench} \quad (12)$$

It should be noted that the $U_{i,j}^{bench}$ correlations between ICSBEP benchmark cases are not quantified yet in the ICSBEP handbook and associated DICE database. However, it is obvious that certain correlations shall exist in the calculated k_i^{cal} results because many measurements were made at the same facilities, with the same fuel rods and/or using the same instrumentation, etc. For example, in OECD/NEA/NSC/WPNCS SG11 analytical exercise, the test calculations were requested for several assumptions on the correlation coefficients due to benchmark uncertainties: equal to 0, 0.7 and 0.99 [10]. In the given study the same values were tested with BM (10),(11), and the results are discussed in section 4.

2.3 Simplified Comparison of the PSI Reference Approach and the Bayesian Solution

It is important to stress that even though the pseudo application cases considered in this work are the canister models with UNF, the approach described in Section 2.1 is the reference PSI methodology for fresh LWR fuel. CSE for burned fuel requires consideration of additional effects and uncertainties (see for details [11]), which, however, are not considered in this work. The focus of this study is to compare the considered methodologies with similar assumptions and inclusion into consideration of the burnup related effects would shift the focus to fuel depletion calculations which goes beyond the scope of this paper.

To facilitate comparison of the different methods, it is convenient to rewrite (1) in the following form with modified USL' option, which becomes application-case specific:

$$k_a^{cal} < \mathrm{USL}' = k_{Normal}^{LTB} - 2\sigma_{\left(k_{Normal}^{LTB} - k_a^{cal}\right)}^{ND} - \Delta k^{AM}. \quad (13)$$

Thus, Eq. (13) corresponds to a modified representation of the PSI reference approach based on k_{Normal}^{LTB}. In certain cases, the non-parametric LTB can be used for verification purposes as discussed in [12] and then the criteria can be expressed as:

$$k_a^{cal} < \mathrm{USL}'' = k_{NP}^{LTB} - 2\sqrt{\sigma_{\left(k_{NP}^{LTB} - k_a^{cal}\right)}^{ND\ 2} + \sigma_l^{\ 2}} - \Delta k^{AM}. \quad (14)$$

In (14), the benchmark specification uncertainty corresponding to the benchmark "*l*" in (7) should be taken into account and it is assumed to be uncorrelated with the nuclear data uncertainty.

In case of the Bayesian solution, an analogous NCS criterion can be considered:

$$k_a^{cal^*} + 2\sigma_{\left(k_a^{cal^*}\right)} < 1 - \Delta k^{AM}, \quad (15)$$

or, equivalently, based on (10):

$$k_a^{cal} < \mathrm{USL}''' = 1 - \Sigma_{ab}U^{-1}\Delta k_i^{bench} - 2\left(\sigma_{\left(k_a^{cal}\right)} - \Sigma_{ab}U^{-1}\Sigma_{ab}^T\right) - \Delta k^{AM}. \quad (16)$$

Note that Eq. (10) provides that if the term $\Sigma_{ab}U^{-1}\Delta k^{bench}$ is negative, then the posterior $k_a^{cal^*}$ is reduced compared to the calculated k_a^{cal}. Equivalently, in (16) it means that the USL′′′ value will be increased. That goes in the less penalizing direction and in the practice of NCS such corrections may not be acceptable [16]. For instance, the American National Standard [17] recommends that in such situations the reasons of the "non-conservative" k_a^{cal} adjustments shall be understood in order to be applied.

In fact, the sign of the term $\Sigma_{ab}U^{-1}\Delta k^{bench}$ is defined by both the signs of individual Δk_i^{bench} biases as well as by the inverse matrix U^{-1} components' signs. This actually makes it complicated to justify the sign of the k_a^{cal} adjustment based on BM, with simple considerations.

For that reason, in the following both cases of acceptance and non-acceptance of a negative $\Sigma_{ab}U^{-1}\Delta k^{bench}$ term will be considered. In the latter case, Eq. (16) will be replaced by Eq. (17):

$$k_a^{cal} < \text{USL}'''' \quad = 1 - 2\left(\sigma_{\left(k_a^{cal}\right)} - \Sigma_{ab}U^{-1}\Sigma_{ab}^{T}\right) - \Delta k^{AM}. \tag{17}$$

Thus, Eqs. (13),(14) and (16),(17) will be used respectively for the comparison of the frequentist and Bayesian based statistics methods presented in Section 4.

3. VERIFICATION STUDIES

3.1 Verification of the PSI NUSS Methodology

Prior to discussing the comparison of the presented above methodologies, it makes sense to outline the verification studies which assist justification of the obtained results. Numerous verifications of the NUSS tool were done in the past [18],[19],[20],[21]. Recently, additional verification was realised using the models considered in this work by comparison with Whisper [22] calculations. The well-known in the field Whisper tool of LANL, working in conjunction with MCNP6® code (see https://mcnp.lanl.gov for details on the MCNP® software trademark), was considered as an appropriate option for verification of the PSI NUSS results, also employing the MCNP6 code for the criticality calculations. However, Whisper uses its own covariance data library (based on the "BLO" library) [22]. For that reason, a direct quantitative comparison was not possible. Nevertheless, a qualitative comparison with the NUSS results could be done, as illustrated below. In both, Whisper and NUSS calculations, the same ENDF/B-VII.1 NDL was used for the neutron transport calculations. The 44-groups energy structure was used with NUSS, similar to a recent study done at PSI [4]. Figure 1 shows the lower triangular part of the correlation matrix for the ICSBEP benchmark cases from the PSI validation suit and the disposal canister models loaded with a PWR UNF with different burnups and initial ^{235}U enrichment about 5 wt%. In the following, the labels "Can-0 ÷ Can-72" correspond to the burnups from 0 to 72 GWd/tHM. The selected burnup values are close to the end of cycle burnups for the ~5wt% enriched fuel irradiated in a Swiss PWR. The burned fuel composition included the major and minor actinides (MA) and fission products (FP), in line with [11],[23]. Cases 'LCT' and 'MCT' are used for illustration and correspond to the UO_2 and MOX ICSBEP benchmarks respectively. The ICSBEP experimental benchmarks with fresh UO_2 are marked as ENNcMM and with MOX fuel as ENNmMM cases in Figure 1.

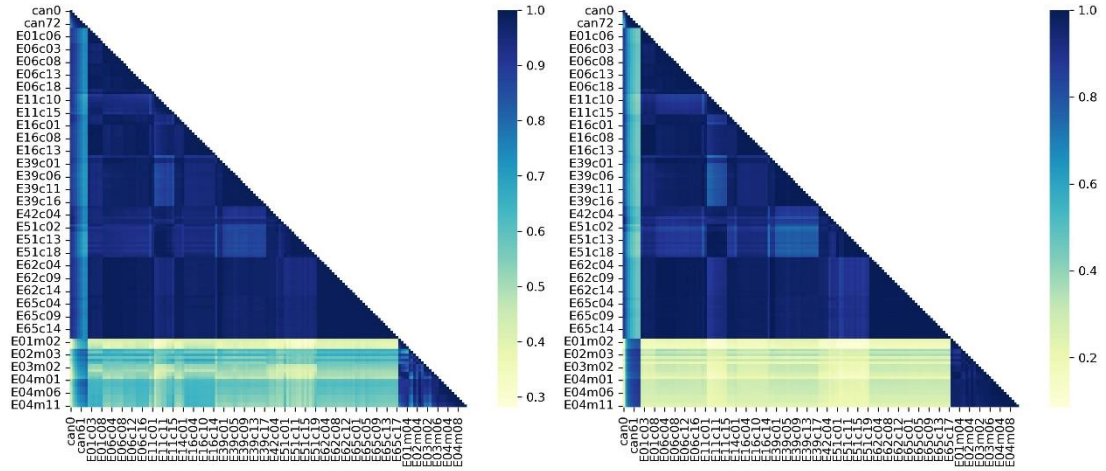

Figure 1: ND-related correlation matrices obtained with NUSS (left) and Whisper (right) for the selected test cases and validation benchmarks.

In the NUSS case, 300 sets of randomly sampled ACE data files were used in addition to the reference NDL. It should be noted that in the NUSS calculations the total NUBAR values (MT-452 in ENDF-6 format) were selected for this study (resulting in more penalizing assessments), although this choice can be questionable [18] and in another similar calculations it was decided to use only the prompt NUBAR values (MT-456)[3]. It should also be mentioned that in the previous works [23],[24] NUBAR and CHI data were not perturbed, while this is the case in the given study and also in [4]. Furthermore, following some preliminary sensitivity and optimisation studies, the target Monte Carlo precision for the MCNP6 calculations with random ACE files was chosen 50 pcm, as reported in [7] (note that more precise MCNP calculations with ENDF/B-VII.1 were reported in [12]). Even though the NUSS and Whisper results show noticeable differences, qualitatively they look rather similar and this justifies that both methodologies provide reasonably consistent results (noting that the differences clearly should be attributed to the different sources of the ND covariance data). Figure 2 represents a zoomed view of the part of the correlation matrix of Figure 1 as concerns the canister models pseudo-application cases, together with a few representative ICSBEP benchmarks. The case E51c16 provides the value in the C/E sample corresponding to the order statistics 95%/95% k_{NP}^{LTB}.

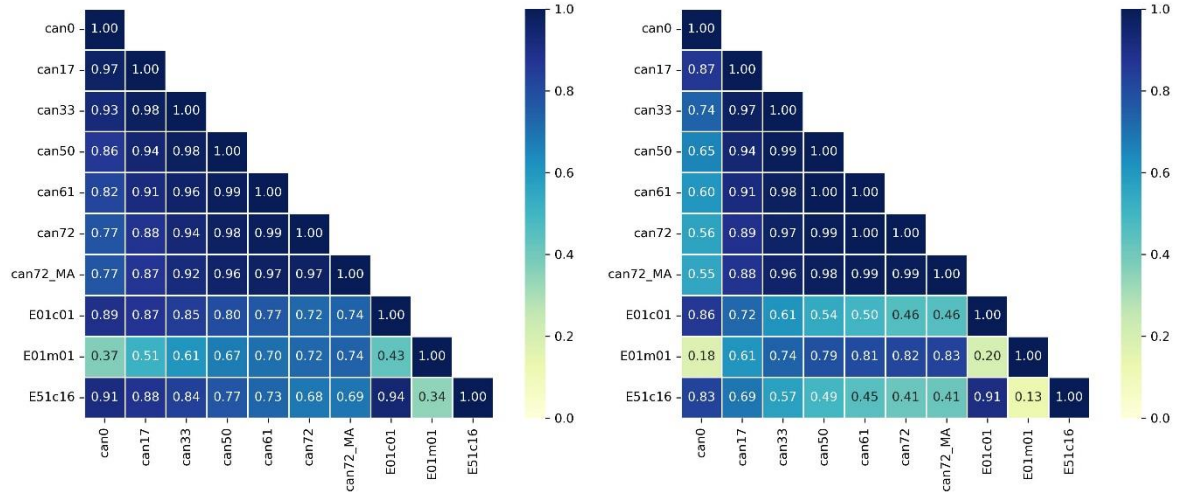

Figure 2: ND-related correlation matrices obtained with NUSS (left) and Whisper (right) for the selected test cases and representative benchmarks.

As it should be expected, with the increase of the burnup the correlations between the canister models and the UO_2 benchmarks are decreasing, while the situation is opposite for MOX fuel.

3.2 Verification of the Bayesian Methodology Realisation

The OECD/NEA/NSC/WPNCS SG11 exercise "Bias and correlated Data, Comparison of Methods" was proposed in 2021 for comparison between methodologies to compute k_{eff} predictions based on given covariance information for the experiments taken into account in the validation procedure [10]. This exercise allowed participants to verify their calculation procedures and tools which could be of relevance to CSE, although not necessary fully consistent with the participants' reference methodologies for practical applications. In particular, PSI submitted two types of solutions for the SG11 exercise, though not fully in line with the reference PSI CSE methodology described in Section 2.1.

Figure 3 illustrates the PSI solutions submitted for the SG11 comparative study. The PSI 'MOCABA' solutions were perfectly matching analogous Bayesian-based solutions and in particular the MOCABA ones provided by other participants [10]. This benchmarking experience confirms that the realisation of the MOCABA methodology, as well as the Bayesian deterministic method employed by PSI (in comparison with the Monte Carlo based MOCABA procedure) [15], corresponding to Eqs. (10),(11) in this study, was done correctly.

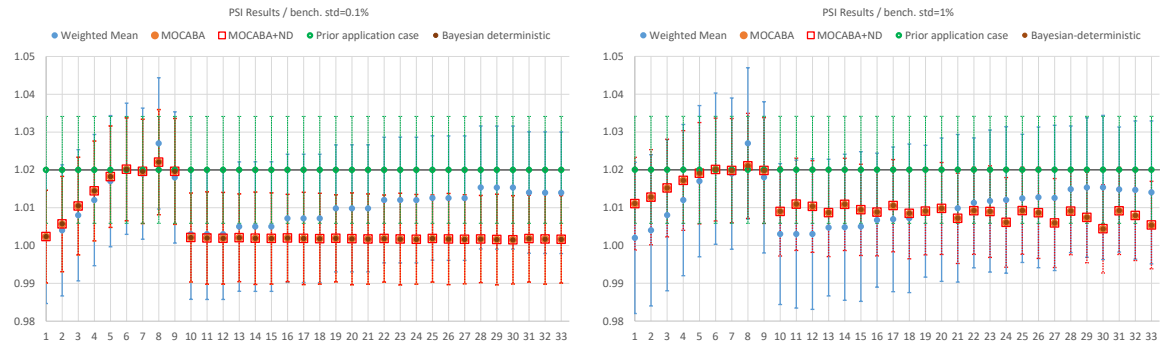

Figure 3: PSI solutions for the WPNCS SG11 benchmark.

The solution marked as "Weighted Mean" [10] was obtained with a modified version of Eq. (4) with inclusion of the full covariance matrices in the mean value derivation instead of only the variances as shown with Eqs. (4),(5). This solution is also called "Covariance Weighted Mean" (CWM) in [7]. Details of that solution go beyond the scope of this paper and will be presented in the WPNCS SG11 final report under preparation. Nevertheless, it is relevant to notice from the PSI results that when only the validation benchmarks with high correlations with an application case are selected and the benchmark uncertainties are low (cases 10-18 in Figure 3), then the "Weighted Mean" solution is close to the Bayesian/MOCABA ones. This is not the case if low-correlated benchmarks are included into the CWM solution (cases 19-33), while in the case of Bayesian/MOCABA the low-correlated benchmarks play no role and the solution is driven almost solely by the higher correlated cases (case 1 in Figure 3 is based on the single most correlated benchmark). In general, this observation means that the PSI reference solution described in Section 2.1 tends to agree better with the Bayesian based solution if only benchmarks highly correlated to an application case are selected for C/E analysis (in other words, if a cut-off value for the correlation coefficient would be applied). This observation is fully in line with the conventional practice to select validation benchmarks close by their physical characteristics to an application case.

4. CALCULATION RESULTS

The relevant calculation results to be presented here for the following discussion are at first the calculation biases for the benchmark cases, Δk_i^{bench} from Eq. (10), illustrated in Figure 4. It should be recalled that for the present study with the "costly" ND uncertainties quantifications, a reduced MC precision was used and for that reason the obtained results are not exactly the same as reported in the previous works, e.g. [12]. Just for the sake of better illustration of the validation results, it can be noted that a simple average of the absolute bias Δk_i^{bench} used in the Bayesian model framework is very close to zero and equals to +1.5pcm.

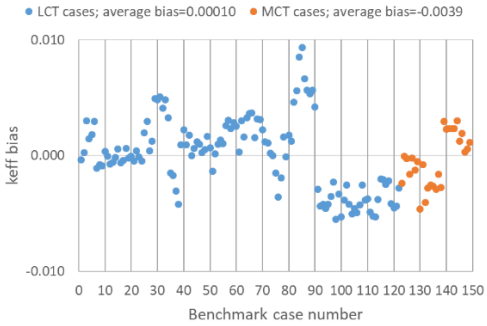

Figure 4: Absolute calculation biases obtained for the PSI CSE validation suit.

Additional important results obtained with the NUSS tool in conjunction with MCNP6 are the ND-related uncertainties, e.g. $\sigma_{k_a^{cal}}^{ND}$ for the application case and the vector of the Pearson correlation coefficients r between the here considered pseudo application cases and the validation benchmarks, as illustrated in Figure 5.

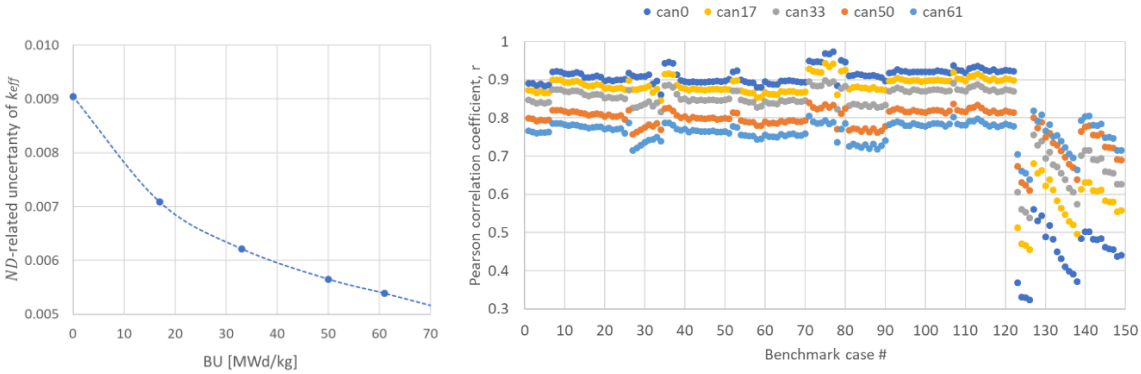

Figure 5: ND-related uncertainties for the application cases (left) and correlations between the application cases and the validation benchmarks (right).

It should be highlighted that the number of NUSS samples (300) and the selected target MC precision (50pcm) lay on rather low sides regarding the statistical convergence of the results. In fact, the standard errors of the Pearson correlation coefficients (see, e.g. [14]) lay in the range from 0.3% to 5.2% for the correlation coefficients respectively in the range from 0.97 to 0.32 shown in Figure 4. This may be considered as rather low statistical precision and should be improved in the future production calculations.

The uncertainty results are obviously dependent on the used NDL and CM. In particular, for the selected case of ENDF/B-VII.1 NDL and CM and the use of the TOTNU CM, it happens that the ND-related uncertainties for the LCT cases (no plutonium in the fuel) are lager compared to the MCT cases (with plutonium in the fuel): ~930pcm vs. ~710pcm on average.

Similarly, one can see in Figure 5 that the ND-related uncertainties are decreasing with burnup (i.e. with decrease of U-235 content and increase of plutonium isotopes content). Naturally, using other sources of CM may lead to different results and trends.

One can notice that the correlation results in general make logical sense: the fresh fuel k_{eff} values are highly correlated with the UO_2 (LCT) benchmarks' k_{eff} values (cases 1-122) while the correlations with the MOX (MCT) benchmarks (cases 123-149) are low. With increasing burnup, the correlations with UO2 benchmarks are decreasing, while the correlations with the MOX benchmarks are increasing. It is also obvious that there are no highly correlated benchmarks for burned fuel. That means that relatively large portions of the k_{eff} calculation uncertainties (can be estimated as 1-r^2) are not shared between the fresh fuel benchmarks and burned fuel application cases. However, it should be mentioned that the use of the "administrative margin" Δk^{AM}, typically equal to 0.05, should reliably cover such sort of validation deficiencies.

Using the obtained results and Eq. (13),(14) and (16),(17) the "modified versions of USL" values can be calculated deterministically for the three considered methods. As was discussed in Section 2.2, the BM equations require knowledge of correlations between k_i^{cal} due to the benchmark model specifications. For the present illustration, the correlation value of 0.7 was arbitrary selected. The respectively obtained results are shown in Figure 6. It is important to note that the considered values are not exactly the USL values as defined by (1), but they are modified and fictitious versions used here only to facilitate the comparison between the different methods.

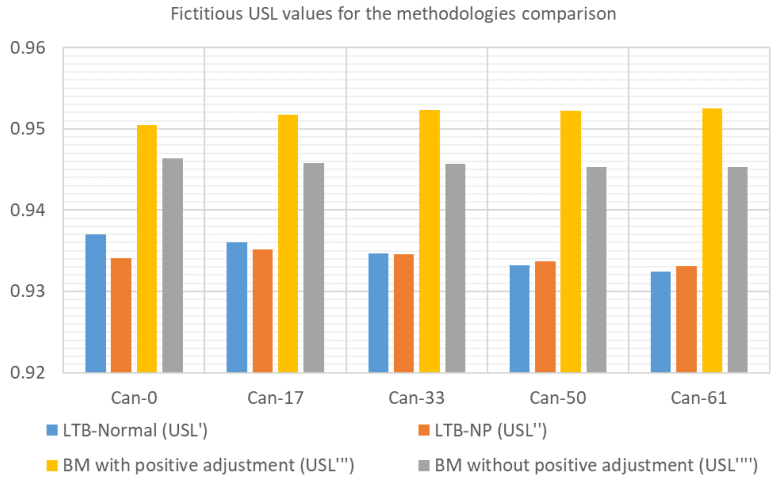

Figure 6: Comparison of the frequentist and the Bayesian –based methodologies performance.

It is further necessary to note that the values of $\Sigma_{ab}U^{-1}\Delta k^{bench}$ discussed in Section 2.3 appeared to be negative for all burnups, meaning that positive adjustments of the USL''' values take place according to the BM equations. It looks counterintuitive that the USL''' values grow with burnup, despite the fact the correlations between the application cases and the benchmarks in general become lower with increase of the burnup, as was shown in Figure 5.

There is a competing effect from the reduction of the application case ND-related uncertainties shown in Figure 4. However, this effect alone does not explain either the negative sign for $\Sigma_{ab}U^{-1}\Delta k^{bench}$ obtained in this work, or the trend with increase of burnup.

It shall be recalled here that an implausible influence of the ND-related CM was observed in [7] in derivation of the weighted mean similarly to Eq. (4), but using the full CM instead of only the variance terms. In the present work, exactly the same matrix was used and the results obtained with the Bayesian model using the ND-related CM for k_a^{cal} adjustment noticeably differ from the frequentist statistics results. Therefore, to verify if the present BM model results may be as well affected by somewhat problematic CM obtained specifically with the employed NDL and CM data (as was discussed in [7]), it is recommended to repeat in the future the present study with alternative sources of ND, e.g. using ENDF/B-VIII or JEFF-3.3.

Ignoring the positive k_a^{cal} adjustment with the $\Sigma_{ab}U^{-1}\Delta k^{bench}$ term when it is negative, i.e. using Eq. (17) instead of (16), leads to slightly more penalising USL'''' values which do not show anymore the counterintuitive growth with burnup. Nonetheless, the BM results even without the positive adjustment are still much less penalising comparing to the frequentist statistics results, meaning that the BM results would be more beneficial considering the economics metrics alone.

The distribution-free or non-parametric LTB method provides slightly more conservative results than the Gaussian-based LTB method for fresh fuel, in line with usual expectations and the previous studies [12]. However, with burnup the difference between the two methods is decreasing and the Gaussian-based LTB method becomes even more penalising in comparison with the non-parametric LTB method. The fictitious USL values become more penalising with burnup for both cases of Gaussian-based LTB method and the Bayesian model method without the positive adjustment. This is related to the Pearson correlation coefficient between the application cases and the fresh fuel validation benchmarks, which is decreasing with growing burnup of the application case models.

As for the assumption on the benchmark k_{eff} correlations, the previously shown BM results without the positive adjustment, obtained for the medium correlations equal to 0.7, were compared with two other options used in the SG11 study: no correlations and strong correlations equal to 0.99. The results are shown in Figure 7.

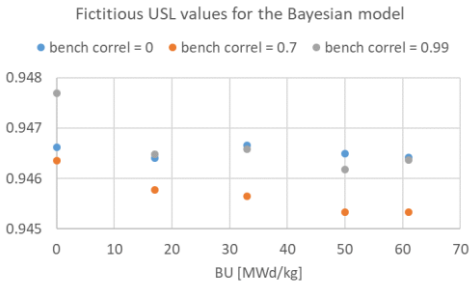

Figure 7: Influence of the benchmark k_{eff} correlations due to the benchmark models/specifications on the Bayesian model results.

One should acknowledge that certain fluctuations of the illustrated results may be related to the low statistical precision of the performed calculations, as was notified above. Nevertheless, Figure 7 indicates that the initially selected value of the benchmarks correlations of 0.7 seem to produce the most penalising results and thus will remain as the reference case for the BM methodology presented here.

For practical BUC for UNF handling, storage, transportation and disposal, it is relevant to translate the differences in the fictitious criticality safety criteria shown in Figure 6 into the difference in required burnup to guarantee sub-criticality.

Such estimation is illustrated in Figure 8, where the two most relevant versions of USL are shown together with the representative UNF canister model k_{eff} behaviour as function of the fuel burnup. In this illustration, the same results as presented in [5] were used; the burnup of 5wt% initially enriched PWR fuel was simulated with a pin cell model, the fuel composition was translated into a canister model and criticality calculations were performed as function of burnup. All relevant details can be found in [5][6].

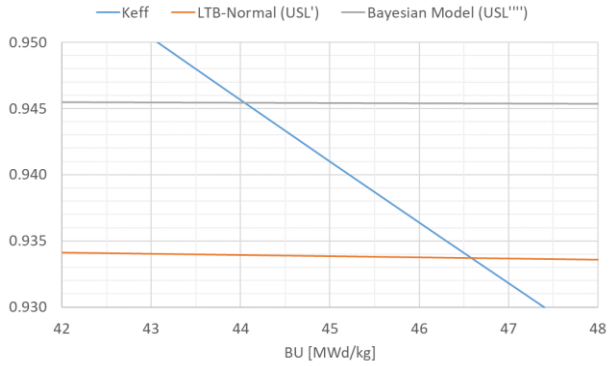

Figure 8: Illustration on the impact of the fictitious USLs' difference for the BUC application.

One can see that the BM approach can lead to some relaxation of the burnup requirement for the UNF criticality safety, namely by ~2.5MWd/kg for the considered case of the 5wt% initially enriched PWR fuel (when using the simplified fuel depletion modelling of [5]).

5. CONCLUSIONS

This work presents a comparison of two intrinsically different options for BUC applications. One is based on the frequentist statistics concept of tolerance bounds and another is based on the Bayesian framework. Both of them have their own advantages and disadvantages. For instance, the frequentist statistics approach utilised in this work does not provide a capability (so far) to include the information about similarity of the validation benchmarks and application cases in the LTB values. All benchmarks are considered equally well representing an application case. Only the ND-related uncertainty component depends on the correlations between the application case and validation benchmarks' k_{eff} values.

The Bayesian model results were found noticeably less penalizing in terms of the considered "fictitious" criticality safety criteria, which, for instance, can result in some practically relevant relaxation of the requirements to the UNF disposal canister designs and/or the UNF loading specifications (i.e. the 'loading curves' [11]). On the other side, the Bayesian model is very sensitive to the C/E results of the benchmarks with highest correlations with an application case, as was for instance illustrated with the SG11 exercise [10]. This potentially can lead to wrong k_{eff} adjustments if the single most important benchmark specifications (i.e. the experimental value "E") or calculation results ("C") are not correct. The most important observation, however, was the counterintuitive behaviour of the safety margins which become less penalising with increasing burnout, although no benchmarks with used nuclear fuel are available in the employed validation suite. Such performance of the utilised methods should be further verified, preferably using different application systems and validation databases.

Nevertheless, in general, the results obtained in this work confirm that both considered methodologies agree reasonably well in terms of the foreseen NCS criteria realisation and the resulting differences in the burnup requirements to guarantee criticality safety are not very large, even if noticeable. Further verification of the obtained findings should be also done with more realistic depletion calculations and improved statistics. As well, latest NDLs and related CMs should be tested. Actually, the present work goes in line with the PSI contribution to the OECD/NEA/NSC/WPNCS SG13 exercise [6], which was designed to foster a consistent comparison of different methodologies using the simplified pseudo-application case with PWR burned fuel. Thus, additional relevant results will be presented in the SG13 final report.

ACKNOWLEDGMENTS

This work was realized partly in relation to the PSI internal project "CROSS2019", funded by the NES and LOG PSI divisions, and also with partial support from Nagra through the joint PSI-Nagra research project BUCSS-R. The authors are grateful to Dr. Madalina Wittel, Nagra, for valuable comments provided to this paper.

6. REFERENCES

[1] B.S. Everitt, A. Skrondal, "The Cambridge dictionary of statistics", Cambridge, New York, Cambridge University Press, 4th ed., 2010.

[2] J.C. Neuber, A. Hoefer, "Frequentist and Bayesian approach in criticality safety uncertainty evaluations", Proceedings of the Topical Meeting of the ANS Nuclear Criticality Safety Division (NCSD 2009), Richland, Washington, USA, 2009.

[3] A. Vasiliev, D. Rochman, M. Pecchia, H. Ferroukhi, "On the options for incorporating nuclear data uncertainties in criticality safety assessments for LWR fuel", *Ann. Nucl. Energy*, 2018, vol. 116, pp. 57-78.

[4] M. Frankl, M. Hursin, D. Rochman, A. Vasiliev, H. Ferroukhi, "Nuclear data uncertainty quantification in criticality safety evaluations for spent nuclear fuel geological disposal", *Applied Sciences,* 2021, vol. 11, 6499 pp. 1-19.

[5] A. Vasiliev, M. Frankl, D. Rochman, and M. Wittel, "Towards a direct comparison of practical CSE with BUC approaches: benchmark proposal for a pseudo-application case with user-defined NCS criteria", Proceedings of ICNC 2023 - the 12th Int. Conf. on Nuclear Criticality Safety, Sendai, Japan, October 1–6, 2023.

[6] A. Vasiliev, M. Frankl, D. Rochman, M. Wittel, A. Hoefer, "Comparison of criticality safety evaluations + burn up credit approaches: assessment of a pseudo-application case with user-defined NCS criteria (WPNCS SG13)", Nuclear Energy Agency; NEA/NSC/WPNCS/WD(2023)01.

[7] A. Vasiliev, H. Lee, M. Frankl, M. Hursin, H. Ferroukhi, "On the PSI routine criticality safety evaluation methodology and its validation approach", Proceedings of ICNC 2023 - the 12th Int. Conf. on Nuclear Criticality Safety, Sendai, Japan, October 1–6, 2023.

[8] S. Chiba, and D.L. Smith, "Impacts of data transformations on least-squares solutions and their significance in data analysis and evaluation", *J. Nucl. Sci*. 1994, vol. 31:8, pp. 770-781.

[9] R. Frühwirth, D. Neudecker, and H. Leeb, "Peelle's Pertinent Puzzle and its solution", *EPJ Web of Conferences*, 2012, vol. 27, Article ID 00008, pp. 1-6.

[10] A. Hoefer et al., "Bias and correlated data, comparison of methods", Proceedings of ICNC 2023 - the 12th Int. Conf. on Nuclear Criticality Safety, Sendai, Japan, October 1–6, 2023.

[11] M. Frankl, A. Vasiliev, D. Rochman, H. Ferroukhi, M. Wittel, and S. Pudollek, "Refinement of the loading curve determination methodology and modeling for Swiss PWR spent fuel final disposal canisters", Proceedings of ICNC 2023 - the 12th Int. Conf. on Nuclear Criticality Safety, Sendai, Japan, October 1–6, 2023.

[12] H. Lee, A. Vasiliev, H. Ferroukhi, "Addressing practical aspects of the nuclear criticality safety evaluations with frequentist statistics", *Ann. Nucl. Energy*, 2023, vol. 181, 109556, 17p.

[13] T. Zhu, A. Vasiliev, H. Ferroukhi, A. Pautz, 2015, "NUSS: A tool for propagating multigroup ND covariances in pointwise ACE-formatted ND using stochastic sampling method", *Ann. Nucl. Energy* 2015, vol. 75, 713-722.

[14] A. Vasiliev, D. Rochman, M. Pecchia, H. Ferroukhi, "Exploring stochastic sampling in nuclear data uncertainties assessment for reactor physics applications and validation studies", *Energies* 2016, vol. 9, 1039.

[15] A. Hoefer, O. Buss, M. Hennebach, M. Schmid, D. Porsch, "MOCABA: A general Monte Carlo–Bayes procedure for improved predictions of integral functions of nuclear data", *Ann. Nucl. Energy*, 2015, vol. 77, pp. 514–521.

[16] US NRC, 2001, "Guide for validation of nuclear criticality safety calculational methodology", NUREG/CR-6698. US Nuclear Regulatory Commission, Washington, DC.

[17] ANS, 2007, "Validation of neutron transport methods for nuclear criticality safety calculations", ANSI/ANS-8.24-2007. American Nuclear Society, IL 60525 USA.

[18] D. Rochman, A. Vasiliev, H. Ferroukhi, T. Zhu, S.C. van der Marck, A.J. Koning, "Nuclear data uncertainty for criticality-safety: Monte Carlo vs. linear perturbation", *Ann. Nucl. Energy*, 2016, vol. 92, pp. 150-160.

[19] W. Wieselquist, T. Zhu, A. Vasiliev, H. Ferroukhi, "PSI Methodologies for ND uncertainty propagation with CASMO-5M and MCNPX: Results for OECD/NEA UAM benchmark phase I", *Sci. Technol. Nucl. Install*, Article ID 549793.

[20] Y. Zhu, D. Rochman, A. Vasiliev, H. Ferroukhi, W. Wieselquist, A. Pautz, "Comparison of two approaches for ND uncertainty propagation in MCNPX for selected fast spectrum critical benchmarks", *ND Sheets*, vol. 118, pp. 388-391.

[21] T. Zhu, A. Vasiliev, H. Ferroukhi, D. Rochman, A. Pautz, "Testing the sampling-based NUSS-RF tool for the nuclear data–related global sensitivity analysis with Monte Carlo neutronics calculations", *Nucl. Sc. and Eng.*, vol. 184, pp. 69-83.

[22] F.B. Brown, M.E Rising, J.L. Alwin, "User manual for Whisper-1.1", LA-UR-17-20567, 2017, 15p.

[23] A. Vasiliev, J. Herrero, D. Rochman, M. Pecchia, H. Ferroukhi, S. Caruso, "Criticality safety evaluations for the concept of Swiss PWR spent fuel geological repository", Proceedings of BEPU 2018, Real Collegio, Lucca, Italy, May 13-19, 2018.

[24] J.J. Herrero, D. Rochman, O. Leray, A. Vasiliev, M. Pecchia, H. Ferroukhi, S. Caruso, "Impact of nuclear data uncertainty on safety calculations for spent nuclear fuel geological disposal", *EPJ Web of Conferences* 2017, vol. 146, Article ID 09028.