

Monte Carlo integral adjustment of nuclear data libraries – experimental covariances and inconsistent data

Henrik Sjöstrand^{1,*} and Georg Schnabel¹

¹Uppsala University, Department of Physics and Astronomy, Box 516, 751 20 Uppsala, Sweden

Abstract. Integral experiments can be used to adjust nuclear data libraries. Here a Bayesian Monte Carlo method based on assigning weights to the different random files is used. If the experiments are inconsistent within them-self or with the nuclear data it is shown that the adjustment procedure can lead to undesirable results. Therefore, a technique to treat inconsistent data is presented. The technique is based on the optimization of the marginal likelihood which is approximated by a sample of model calculations. The sources to the inconsistencies are discussed and the importance to consider correlation between the different experiments is emphasized. It is found that the technique can address inconsistencies in a desirable way.

1 Introduction

Integral experiments (benchmarks) can be used to adjust nuclear data (ND) libraries. The process is referred to as Integral adjustment (IA). The motivation is to use the full width of the experimental data available and hence increase the predictive power of the ND library. IA is usually done implicitly, when models, parameters, and experiments are chosen in order to assure proper performance of the ND library in respect to the integral data. Alternatively, the IA can be performed explicitly, where the ND-library is tuned to the integral data using fitting routines and mathematical models. This paper relates to such explicit IA, and from here on the use of IA will only refer to explicit IA.

IA is a well-established field [1, 2] and can be performed by deterministic adjustment techniques [1] or by Monte Carlo IA (MC-IA) [3, 4]; this paper concerns the later which sometimes is referred to as Bayesian Monte Carlo (BMC) [5]. However, since BMC can be used for both differential- and integral data and this particular paper addresses the use of integral data, we will use the term MC-IA

MC-IA has been explored by many authors, both for improving the best estimate (central value) of the ND file [6], but also to reduce the ND uncertainty [2 - 4]. In this paper, the MC-

* Corresponding author: henrik.sjostrand@physics.uu.se

IA method is extended by including a marginal likelihood optimization (MLO) technique to address inconsistencies between the experimental and calculated integral parameter (in this case k_{eff}). The paper also discusses the sources of these inconsistencies. Methods for treating inconsistent integral data has been proposed earlier, such as the adjustment margin (AM) method [1], and the $\Delta\chi^2$ -filtering [7]. The AM method does not consider correlations and is binary. The $\Delta\chi^2$ is also binary, and the choice of the rejection criteria appears to be rather arbitrary. In a binary method, an experiment is either kept or rejected. I.e., it is assumed that we have either full or no confidence in the experiment. We believe that this is a too coarse selection mechanism and that a continuous scale representing our trust in an experiment is more appropriate. The proposed method is statistically well-founded, and address the mentioned issues of the previous methods. The treatment of inconsistent data and handling of outliers is a wide research field, and the above-mentioned methods are examples of techniques that have been used in the IA community. Many other methods have been suggested for treatment of inconsistent data and Ref. [8] compares the performance of some different methods. In [8], the Bayesian procedure to adjust uncertainties performed best. The presented Bayesian procedure determines a factor to rescale the uncertainties of all data points. In contrast, our approach allows uncertainties to be adjusted individually. In the Bayesian setting, our MLO approach corresponds to the maximum a posteriori probability estimate of the uncertainties.

This paper is outlined in the following way, in chapter two the basics of MC-IA is explained as well as the underlying motivation for adding a calculation uncertainty when using IA. An example of an MC-IA with inconsistent data is also provided. Chapter three describes the MLO technique. Chapter 4 presents the results when using the MLO technique in combination with the MC-IA. Chapter 5 contains the conclusions.

2. Monte Carlo Integral adjustment

In IA, the ND is updated to reduce the deviation between the calculated and experimental value of the integral experiment. There are many different types of integral experiments and in this paper criticality experiments from the ICSBEP [9] is used. More specifically, the ICSBEP MCNP benchmarks of the integral experiments are used to calculate the k_{eff} of the benchmarks. This is subsequently compared to the reported benchmark values. A reported benchmark value is based on the experimental result of the underlying integral experiment and hence referred to as the experimental value.

MC-IA uses ND in the form of random files. Each ND random file is a full ENDF (but without a covariance file), which can be a possible realization of the underlying ND given the PDF of the ND. I.e., the set (e.g., 1000) of ND random-files is a way to provide the PDF of the ND. ND-random files can be obtained through the so-called Total Monte Carlo (TMC) method [10, 11]. These TMC random files are provided on the TENDL webpage or can be produced using the so-called T6 code package [10]. Alternatively, ND random files can be generated by sampling the covariance matrix of an ENDF (e.g., JEFF3.3). This sampling can be achieved by using tools like SANDY [12, 13], NUSS [14], or NUDUNA [15]. One advantage of using TMC random files is that these can contain higher order moments of the ND-PDF as well as more cross-correlations between different parts of the ND file.

Given a reasonably large set of random-files (at least 1000, but in many cases more files are needed for convergence), an MC-IA can be performed. Each random file is used in the benchmark calculations, e.g., calculating the k_{eff} for a set of benchmarks. The MC-IA is performed by assigning weights, w , to the different random files, reflecting the agreement between the calculated and the experimental value. The weight is set to,

$$w_i = \delta e^{-\frac{\chi_i^2}{2}}, \quad (1)$$

(see, e.g., [16 - 18]) where w_i is the weight for random file i , δ is an arbitrary normalization constant and χ_i^2 , is defined as

$$\chi_i^2 = \sum_b \frac{(C_{b,i} - E_b)^2}{\sigma_b^2}, \quad (2)$$

where b is the benchmark number, $C_{b,i}$ is the calculation result of benchmark b for random file i , E_b , is the experimental value of benchmark b and σ_b is the benchmark uncertainty. If correlations are also considered we instead use the generalized χ^2

$$\chi_i^2 = (\mathbf{C}_i - \mathbf{E}_i)^T \mathbf{B}^{-1} (\mathbf{C}_i - \mathbf{E}_i), \quad (3)$$

where \mathbf{C}_i and \mathbf{E}_i are the vectors with the calculation and experimental results for all benchmarks. \mathbf{B} is the covariance matrix for the benchmarks. \mathbf{B} or σ_b should contain the reported experimental uncertainty σ_{reported} , the statistical uncertainty of the MCNP calculations, σ_{stat} , and possibly other calculational uncertainties [4] (see Eq. 4 for details). In Refs. [19, 20] the same approach is used, but in [20] with a different definition for the weights.

When the weights for the random files have been calculated, a new PDF for the underlying ND can be determined [3]. In addition, a new weighted PDF for the calculated value of the benchmark or for any other macroscopic parameter for any system under investigation can be inferred. I.e., using the weights, a posterior best estimate and ND uncertainty can be deduced for, e.g., the benchmarks used in the IA. This is also how the weights are used in this work. An example of this is seen in Fig. 1 where 1000 TENDL 2014 random files for ^{235}U and ^{238}U are used to calculate k_{eff} for five benchmarks. These benchmarks and random files are used for all the results in the paper.

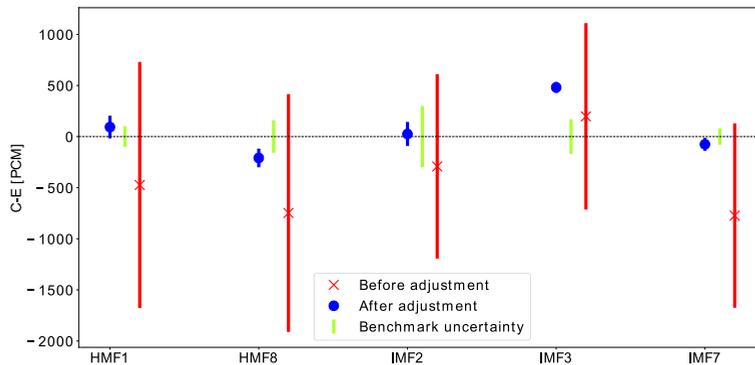


Fig. 1. (Color online) The figure shows the results for five criticality benchmarks. On the y-axis is the difference between the mean of $C_{b,i}$ (over i) and the experimental value in PCM. The red crosses are the results before the weighting of the random files. The blue dots are the result after the weighting of the random files using Eq. 1. The error-bars connected to the crosses and dots contain the ND uncertainty of the calculated k_{eff} . The bars without marks are the benchmark uncertainties.

As can be seen in Fig. 1, all except IMF3 obtain better agreement (in absolute values) after the IA, and the ND uncertainty is reduced for all the benchmarks. In fact, the ND uncertainty is reduced to such an extent that the calculation results are hardly consistent with the experimental results. A p-value of 6 % is obtained if both experimental and ND uncertainty is included. The p-value is defined as the probability that the difference between the experiments and the calculation would be at least as extreme as the obtained result, given

the null hypothesis. In our case, the null hypothesis is that the difference between the experimental data and the model prediction is governed by a multivariate normal distribution with a zero mean-vector and a covariance matrix composed as the sum of the experimental covariance matrix and the model covariance matrix. The test statistic is the generalized chi-square value and "as least as extreme" has to be understood with respect to this test statistic. IMF3 cannot be adjusted to obtain a good agreement with the experimental value simultaneously as the other values were adjusted because of the strong calculation correlations between the benchmarks (see Fig. 2 left). E.g., if the fission cross-section in the fast energy-range is increased for ^{235}U , k_{eff} is increased for all the benchmarks.

In fact, by eye one can easily believe that the prior data (data before the adjustment) is in good agreement with the experimental data. However, if the p-value is calculated for the prior data including the full ND covariance a p-value of 11% is obtained. The reason for this low p-value is the strong correlations between the benchmarks. What we witness is that the calculation data disagrees with our set of experimental data. Differently expressed, given the calculational data there is a disagreement between the experiments. This disagreement can be due to:

- Model defects. I.e., the true ND is not within the PDF of the reported ND (in our case reported using random files).
- The adjusted ND (in this case ^{235}U and ^{238}U) contain errors that are not reported in the covariance file or represented by the random files. E.g., in many libraries, covariance information for angular distribution is missing.
- Unaccounted experimental uncertainties or covariances.
- Isotopes not considered. In this case, only ^{235}U and ^{238}U are adjusted. However, the actual experiments contain other isotopes that could be responsible for the disagreement. These isotopes can either be included in the adjustment procedure or should be marginalized, i.e., added to the benchmark uncertainty [4].

Whatever the reason, it is clear that the posterior uncertainties are too small given our set of experiments. This disagreement needs to be addressed in some way.

3 Addressing inconsistent experiments

As mentioned previously there are many reasons for the data to be inconsistent, both among the different experiments and between the experiments and the calculation. Many of these underlying causes can be treated by different approaches. E.g., model defects can be considered in new ND evaluations [21, 22] or better models can be used; ND evaluations can include more uncertainties (an example of this is the TENDL evaluations [10], which report uncertainties on most ND quantities); experiments can be re-evaluated to find unaccounted uncertainties; and more isotopes can be considered in the IA [4]. We encourage and are part of these efforts, but also acknowledge that at any given time an IA is performed, remaining inconsistencies need to be treated. In this section, we present the MLO technique for the treatment of such inconsistencies. In [23] a different version of the MLO approach based on the construction of a sensitivity matrix of the model by linearizing it was suggested. Here we do not linearize the model but build the MLO approach on top of a sample of model calculations.

We add individual extra uncertainties, $\sigma_{\text{extra},b}$, to each of the benchmarks to account for the inconsistencies. In addition, we add a fully correlated common uncertainty, $\sigma_{\text{extra},c}$, to all the benchmarks. I.e., for each benchmark the new benchmark uncertainty is calculated as

$$\sigma_b^2 = \sigma_{\text{reported}}^2 + \sigma_{\text{stat}}^2 + \sigma_{\text{extra},b}^2 + \sigma_{\text{extra},c}^2 \quad (4)$$

In this way, we try to model that there can be effects that are common for all the experiments and that there can be additional individual components in each experiment. The

magnitudes of the σ_{extra} are found by using MLO. Details on the technique can be found in [23]. The likelihood for a σ_{extra} is described by,

$$L = \sum_i e^{-\frac{\chi_i^2}{2}} / \sqrt{(2\pi)^n |B|}, \quad (5)$$

where n is the number of experiments and $|B|$ is the determinant of the benchmark covariance matrix, with the components defined in Eq. 4. With increased σ_{extra} , χ^2 decreases and the numerator term increases. However, with increased σ_{extra} , the denominator increases as well. A derivation of Eq. 5 can be found in [24]. An example of the likelihood function in one dimension is shown in Fig. 2 (right), where σ_{extra} for HMF1 is varied and kept fixed for the other experiments.

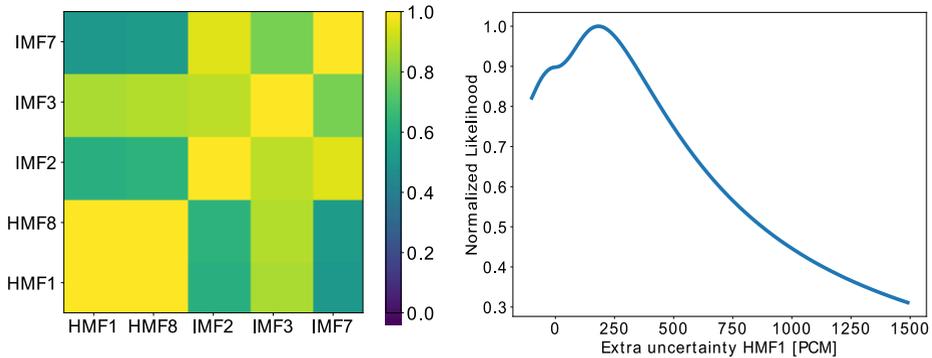


Fig. 2 (Left) The prior calculation correlation between the benchmarks. (Right) The normalized likelihood (y-axis) for the $\sigma_{\text{extra,b}}$ (x-axis) of HMF1.

As can be seen, the most likely σ_{extra} for HMF1 is close to 250 pcm in this example; this value should be used for σ_{extra} . In the full exercise, the six-dimensional space (with five individual $\sigma_{\text{extra,b}}$ and a $\sigma_{\text{extra,c}}$) was explored, and L was optimized (maximized). For the optimization, the minimum of the negative logarithm of L was found using the python `scipy.optimize.minimize` package with the ‘SLSQP’ method. For the calculation of the determinant, a Cholesky decomposition of B was performed using the python `numpy.linalg.cholesky` package.

Given no information from the data, we would generally assume that the σ_{extra} should be small since considerable work has been done in both evaluating the experiments and creating the ND files. Eq. 5. does not consider this prior belief. To promote small σ_{extra} , a new term is added to the likelihood,

$$L = \left(\sum_i e^{-\frac{\chi_i^2}{2}} / \sqrt{(2\pi)^n |B|} \right) e^{-\sum \beta \sigma_{\text{extra}}^2}, \quad (6)$$

where the summation is over both $\sigma_{\text{extra,b}}$ and $\sigma_{\text{extra,c}}$, β is a constant that allows us to shape our prior belief on the σ_{extra} . Here, β is set so that 500 pcm σ_{extra} has a prior probability (before inspecting the data) of 50 % compared to zero σ_{extra} . In the choice of β , expert judgment is introduced in the methodology. We use a Gaussian function in the added term in Eq. 6, but other functions have also been suggested [23]. It is not absolutely necessary to include the term. However, by not including it σ_b becomes independent of σ_{reported} . I.e., the information from the evaluation of the benchmarks uncertainties is not affecting the final results. This would be undesired, and hence we recommend using a prior term. In addition, β allows us to code expert knowledge into the likelihood. In our case β was set to the same value for all the experiments; however, β can be chosen to have individual values for different experiments to encode our belief in a particular experiment.

4 Results

The adjustment was tested including $\sigma_{\text{extra},c}$; without $\sigma_{\text{extra},c}$; using Eq. 5; and using Eq. 6. Fig. 3 illustrates the results using Eq. 6 and including $\sigma_{\text{extra},c}$.

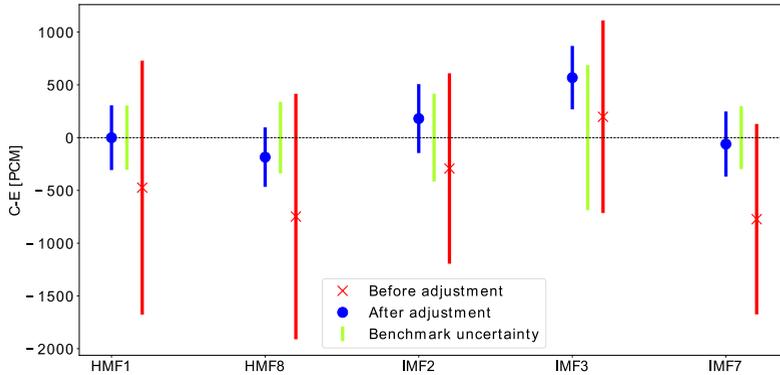


Fig. 3 Same as Fig. 2, but using the MLO technique to include σ_{extra} . The green bars without marker contain the benchmark uncertainty including σ_{extra} .

Fig. 3 now presents a consistent data set. If compared to the results without using the MLO (Fig. 2), one can note that the posterior ND uncertainty is dramatically increased (e.g., from 69 pcm to 253 pcm for HMF1) In this case, the total benchmark uncertainty is augmented for all benchmarks, mostly since a 209 pcm common uncertainty component is added to all the experiments. In fact, only IMF3 obtains a significant individual extra uncertainty added and the $\sigma_{b,IMF3}$ increases from 170 pcm to 468 pcm. The results become quite different if $\sigma_{\text{extra},c}$ is assumed to be zero, as can be seen in Table 1. In the MLO $\sigma_{\text{extra},b}$ were allowed to take negative values as long as the sum of $\sigma_{\text{extra},b}$ and $\sigma_{\text{extra},c}$ was larger than zero. The motivation for this was that part of the reported experimental uncertainties could stem from common causes in the different benchmarks. To assure to retain a minimum individual uncertainty the criterion $\sigma_{\text{extra},b}^2 > -0.5\sigma_{\text{reported},b}^2$ was also included.

Table 1. Table of resulting σ_b . The second row contains σ_{reported} (No MLO performed). In the third row $\sigma_{\text{extra},c}$ is assumed to be zero. The third and fourth row uses Eq.5. The last row uses Eq. 6.

σ_b [PCM]	HMF1	HMF8	IMF2	IMF3	IMF7	$\sigma_{\text{extra},c}$
No MLO	100	160	300	170	80	0
using Eq. 5	153	204	300	580	390	0
using Eq. 5	267	329	333	591	409	257
using Eq. 6	232	263	366	468	228	209

5 Conclusions and outlook

A new technique to address inconsistent integral experiments using an MC-MLO technique is proposed and applied to a set of benchmarks. The method can resolve the inconsistencies in the data, and the outcome of the results is desirable. To our knowledge, the MC version of the technique has not been proposed before in the context of ND-evaluation, and we believe it is the first time an MLO technique is proposed for IA.

The technique has many desirable features, e.g., being statistical well-founded; can be integrated with expert judgment in a transparent way; is not binary; takes into account both experimental and calculational correlations between the benchmarks; and results in

reasonable posterior ND uncertainties. MLO is hence a promising technique to be used for integral adjustment in future library releases.

We suggest that the technique should be tested with a larger set of benchmarks, where also validation benchmarks are used to infer the predictive power of the method. In addition, the method should be compared quantitatively with other discrepant data techniques.

Acknowledgment

This work was supported by the Swedish research council and Swedish Nuclear Centre. A special thanks to Daniel Siefman for input in regards to established discrepant data methods. The benchmark suite used was prepared by Steven Van der Marck.

References

- [1] M. Salvatores, et al., Nuclear Data Sheets **118**, 38 (2014).
- [2] C. De Saint Jean, P. Archier, E. Privas, G. Noguère, O. Litaize, and P. Leconte, Nucl. Data Sheets **123**, 178 (2015).
- [3] D. A. Rochman, E. Bauge, A. Vasiliev, H. Ferroukhi, and G. Perret, EPJ Nucl. Sci. Technol. **4**, 7 (2018).
- [4] E. Alhassan, H. Sjöstrand, P. Helgesson, M. Österlund, S. Pomp, A. J. Koning, and D. Rochman, Prog. Nucl. Energy **88**, 43 (2016).
- [5] C. D. S. Jean, P. Archier, E. Privas, and G. Noguere, EPJ Web Conf. **146**, 02007 (2017).
- [6] D. Rochman and A. J. Koning, Nucl. Sci. Eng. **169**, 68 (2011).
- [7] M. L. Williams, B. L. Broadhead, M. A. Jessee, J. J. Wagschal, and R. Lefebvre, Oak Ridge Natl. Lab. ORNLTM-200539 **3**, (2009).
- [8] S. I. Kafala, T. D. MacMahon, and P. W. Gray, Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip. **339**, 151 (1994).
- [9] *International Criticality Safety Benchmark Evaluation Project (ICSBEP)*.
- [10] A. J. Koning and D. Rochman, Nucl. Data Sheets **113**, 2841 (2012).
- [11] D. Rochman, W. Zwermann, S. C. van der Marck, A. J. Koning, H. Sjöstrand, P. Helgesson, and B. Krzykacz-Hausmann, Nucl. Sci. Eng. **177**, 337 (2014).
- [12] L. Fiorito, G. Žerovnik, A. Stankovskiy, G. Van den Eynde, and P. E. Labeau, Ann. Nucl. Energy **101**, 359 (2017).
- [13] L. Fiorito, *Sandy* (<https://luca-fiorito-11.github.io/sandy-docs/>).
- [14] T. Zhu, A. Vasiliev, H. Ferroukhi, and A. Pautz, Ann. Nucl. Energy **75**, 713 (2015).
- [15] A. Hoefler, O. Buss, M. Hennebach, M. Schmid, and D. Porsch, Ann. Nucl. Energy **77**, 514 (2015).
- [16] R. Capote, D. L. Smith, A. Trkov, and M. Meghzifene, J. ASTM Int. **9**, 1 (2012).
- [17] P. Helgesson, H. Sjöstrand, A. J. Koning, J. Rydén, D. Rochman, E. Alhassan, and S. Pomp, Prog. Nucl. Energy **96**, 76 (2017).
- [18] R. Capote and D. L. Smith, Nucl. Data Sheets **109**, 2768 (2008).
- [19] C. D. S. Jean, P. Archier, E. Privas, and G. Noguere, EPJ Web Conf. **146**, 02007 (2017).
- [20] E. Bauge and P. Dossantos-Uzarralde, J. Korean Phys. Soc. **59**, 1218 (2011).
- [21] G. Schnabel and H. Leeb, EPJ Web Conf. **111**, 09001 (2016).
- [22] P. Helgesson and H. Sjöstrand, Ann. Nucl. Energy **120**, 35 (2018).
- [23] G. Schnabel, ArXiv180300960 Nucl-Ex Physicsnucl-Th Physicsphysics (2018).
- [24] G. Schnabel, (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-381563>).