Nuclear Science
NEA/NSC/R(2021)4
February 2023
www.oecd-nea.org

Investigation of Covariance Data in General Purpose Nuclear Data Libraries

OECD
BETTER POLICIES FOR BETTER LIVES

NEA
NUCLEAR ENERGY AGENCY

Nuclear Energy Agency

NEA/NSC/R(2021)4

Unclassified

English text only

3 February 2023

NUCLEAR ENERGY AGENCY
NUCLEAR SCIENCE COMMITTEE

Investigation of Covariance Data in General Purpose Nuclear Data Libraries

This document is available in PDF format only.

JT03511603

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 38 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Türkiye, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

NUCLEAR ENERGY AGENCY

The OECD Nuclear Energy Agency (NEA) was established on 1 February 1958. Current NEA membership consists of 34 countries: Argentina, Australia, Austria, Belgium, Bulgaria, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, Norway, Poland, Portugal, Romania, Russia (suspended), the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Türkiye, the United Kingdom and the United States. The European Commission and the International Atomic Energy Agency also take part in the work of the Agency.

The mission of the NEA is:

– to assist its member countries in maintaining and further developing, through international co-operation, the scientific, technological and legal bases required for a safe, environmentally sound and economical use of nuclear energy for peaceful purposes;

– to provide authoritative assessments and to forge common understandings on key issues as input to government decisions on nuclear energy policy and to broader OECD analyses in areas such as energy and the sustainable development of low-carbon economies.

Specific areas of competence of the NEA include the safety and regulation of nuclear activities, radioactive waste management and decommissioning, radiological protection, nuclear science, economic and technical analyses of the nuclear fuel cycle, nuclear law and liability, and public information. The NEA Data Bank provides nuclear data and computer program services for participating countries.

Foreword

The understanding of nuclear physics and the associated uncertainties is essential to the modelling of all nuclear systems. Nuclear reaction probabilities, emitted particle energies and angles, fission and other phenomena vary by many orders of magnitude over the energy ranges that neutrons experience within nuclear reactors and the correlations between the uncertainties over these energy ranges are indispensable to the useful estimation of uncertainty for operational parameters.

The Working Party on International Nuclear Data Evaluation Co-operation (WPEC), under the auspices of the Nuclear Energy Agency (NEA) Nuclear Science Committee (NSC), was established in 1989 to promote the exchange of information on nuclear data evaluations, validation and related topics. Its aim is also to provide a framework for co-operative activities among members of the major nuclear data evaluation projects. This framework includes the possible exchange of scientists in order to encourage co-operation. The WPEC determines common criteria for evaluated nuclear data files with a view to assessing and improving the quality and completeness of evaluated data.

This Working Party is organised by the NEA in close co-operation with several parties such as the Russian Evaluated Neutron Data Library (BROND), the Evaluated Nuclear Data File (ENDF, United States), the Japanese Evaluated Nuclear Data Library (JENDL), the Joint Evaluated Fission and Fusion File (JEFF) (with other NEA Data Bank member countries) and the Chinese Evaluated Nuclear Data Library (CENDL) through the Nuclear Data Section of the International Atomic Energy Agency (IAEA).

This report gives an overview of the activities undertaken by WPEC Subgroup 44 (SG-44) on covariances for general-purpose nuclear data libraries. The SG-44 has studied the state of the art in evaluation techniques for nuclear data covariances, methodologies for using integral experiments to generate and update covariance matrices, investigations into general cross-correlations between different isotopes and physics types, and data format extensions to accommodate advanced covariance data. An intercomparison study was also performed to draw conclusions on covariance generation and explore the possibility of application-independent covariance data. This report summarises those studies and draws conclusions for the future evaluation of covariance data for general-purpose data libraries.

Acknowledgements

The NEA wishes to express its sincere gratitude to Professor Vladimir Sobes (United States), who co-ordinated this work and the report, Dr Cyrille de Saint Jean (France), who served as the activity monitor, and to the members of the Subgroup 44 for their contribution to this report.

Table of contents

List of figures

List of tables

List of abbreviations

BMC	Bayesian Monte Carlo
BNL	Brookhaven National Laboratory (United States)
BROND	Russian Evaluated Neutron Data Library
C/E	Calculation and experimental values
CEA	Commissariat à l'énergie atomique et aux énergies alternatives (France)
CENDL	Chinese Evaluated Nuclear Data Library
CIELO	Collaborative International Evaluated Library Organisation
CONRAD	Nuclear data modelling and evaluation code (CEA, France)
CSEWG	Cross Section Evaluation Working Group (US DOE, United States)
DICE	Database for ICSBEP (NEA)
EFF	European Fusion File (NEA)
ENDF	Evaluated Nuclear Data File (United States)
EXFOR	Exchange FORmat
G(L)LS	Generalised (linear) least-squares
GNDS	Generalised Nuclear Data Structure
IAEA	International Atomic Energy Agency
ICSBEP	International Criticality Safety Benchmark Evaluation Project (NEA)
IMF	Intermediate Enriched Uranium Metal Fast-spectrum Benchmarks (ICSBEP, NEA)
IRPhE	International Reactor Physics Evaluation Project (NEA)
JEFF	Joint Evaluated Fission and Fusion File
JENDL	Japanese Evaluated Nuclear Data Library
KENO	Monte Carlo radiation transport code within the SCALE package (ORNL, United States)
LANL	Los Alamos National Laboratory (United States)
LCT	Low-enriched Uranium Compound Thermal-spectrum Benchmarks (ICSBEP, NEA)
MCNP	Monte Carlo N-Particle
MLO	Marginal likelihood optimisation
ND	Nuclear data
NEA	Nuclear Energy Agency
NJOY	Nuclear data processing code (LANL, United States)
NSC	Nuclear Science Committee (NEA)

OECD	Organisation for Economic Co-operation and Development
ORNL	Oak Ridge National Laboratory (United States)
PANTHER	Reactor simulation code (United Kingdom)
PDF	Probability distribution function
PFNS	Prompt fission neutron spectra
PMF	Plutonium Metal Fast-spectrum Benchmarks (ICSBEP, NEA)
PUB	Physical Uncertainty Bounds
PWR	Pressurised water reactor
REFIT	Nuclear resonance parameter analysis code (United Kingdom)
SAMMY	Nuclear resonance parameter analysis code (ORNL, United States)
SANDY	Nuclear data covariance sampling tool (SCK•CEN, Belgium)
SCALE	Modelling and simulation suite for nuclear analysis (ORNL, United States)
SCK CEN	Studiecentrum voor Kernenergie • Centre d'Étude de l'énergie Nucléaire (Belgium)
SUSD3D	3D deterministic sensitivity calculation tool (Slovenia)
TALYS	Nuclear data evaluation code (European Union)
TASMAN	Nuclear data model parameter sampling code (European Union)
TENDL	TALYS-based Evaluated Nuclear Data Library (European Union)
TMC	Total Monte Carlo
TSL	Thermal scattering law
TSURFER	Tool for Sensitivity and Uncertainty Analysis of Response Functions Using Experimental Results (ORNL, United States)
UACSA	Uncertainty Analysis for Criticality Safety Assessment
UMC	Unified Monte Carlo
UQ	Uncertainty quantification
USU	Unrecognised Source of Uncertainties
WPEC	Working Party on International Nuclear Data Evaluation Co-operation (NEA)
WPNCS	Working Party on Nuclear Criticality Safety (NEA)

Executive summary

Knowledge of basic nuclear physics is fundamental to our ability to model, simulate and understand the operation of nuclear systems. State-of-the-art databases in this field contain uncertainties that reflect those of the experiments they are based on. These uncertainties are essential for characterising the level of certainty the community places on simulation results, but care must be taken to ensure these are not overly or unrealistically conservative. Laws of nuclear physics and inferences from integral experiments provide many correlations between physical properties that constrain the variances in nuclear data. When including the correlation terms, these are known as covariances. Covariances have been the subject of intense research over the past decade, with all major nuclear data libraries aiming to provide complete and well-founded covariances for all data they release.

Due to the nature of experimental inference, covariances remain in some senses application specific, and the Nuclear Energy Agency (NEA) Working Party on International Nuclear Data Evaluation Co-operation (WPEC) launched its 44th Subgroup to review the state of the art in nuclear data covariance generation for general-purpose nuclear data libraries. The present report summarises their findings on the full range of input components to covariances, including the current methods for differential nuclear physics data evaluation, selection and application of integral experiments, creation of cross-physics and even cross-isotope correlations, recommendations for data format extensions and the results of an intercomparison study of uncertainty correlations on challenge problems agreed by the subgroup participants. The primary recommendations include the need for detailed documentation and/or reproducible processes for the creation of covariances to ensure users and those that benefit from their analyses are fully informed of the applicability and limitations of simulation results. Specific recommendations were made to the NEA Expert Group on the Generalised Nuclear Data Structure (GNDS), which adopted these in version 2.0 of the format specifications published by the NEA.

1. Introduction

The objective of the subgroup is to bring together the international covariance community to understand how the covariance data can be so different between the different evaluated nuclear data files (ENDF, JEFF, JENDL, CENDL, etc.) while the mean values (cross-sections, nu-bar, etc.) are generally very similar. Many questions have emerged from the groups applying covariance data for analysis, such as the NEA Working Party on Nuclear Criticality Safety (WPNCS) Expert Group on Uncertainty Analysis for Criticality Safety Assessment (UACSA), on how the use of different covariance libraries (e.g. ENDF, JEFF, JENDL) affects uncertainty quantification and similarity assessment. Further, significant differences in covariance libraries lead to differences in the adjustment of parameters for fast reactors, which is an important topic for the NEA WPEC Subgroup (SG-39).

The Collaborative International Evaluated Library Organisation (CIELO) Pilot Project, WPEC SG-40, established an international effort of nuclear data evaluators from different nuclear data projects to provide nuclear data evaluations that may be consistently accepted by all major nuclear data projects. This work certainly drove progress in minimising disagreement in the mean values (cross-sections, nu-bar, etc.) between nuclear data libraries. However, as that project came to a close, there had not been a concentrated effort to provide consistent covariance evaluations across nuclear data libraries. The maturity of the nuclear data evaluation process was such, at the time, that an international collaboration on cross-section covariance evaluation methodologies was warranted.

This subgroup was tasked with the goal of investigating covariance data for a broad range of system types, not just fast reactors, as was the focus of WPEC SG-39. This subgroup leveraged the work of previous subgroups that had investigated the generation of covariance data for specific physical regions, including WPEC SG-24 and SG-36, which had focused on evaluations of the fast neutron region and the resolved resonance region, as well as WPEC SG-42, which had focused on the evaluation and covariance generation for thermal scattering. This subgroup focused its attention on providing guidance to the international community on methods for systematic and consistent evaluation of covariance data for the whole energy range, paying special attention to energy domain interface (resolved resonance/unresolved resonance/continuum). The group also delivered examples of the application of the proposed methodology on a few selected isotopes. The ultimate goal of the subgroup was to provide an overview of best practices to generate more consistent covariance data sets.

2. Main evaluation and techniques

2.1. Overview of the methodologies

Since the beginning of nuclear physics and related applications, having nuclear data based on available experiments complemented with knowledge of theoretical nuclear physics was mandatory. The evaluation of these data is still based on the same philosophy: a synthesis of knowledge coming from nuclear reaction models, differential measurements as well as integral measurements.

The main evaluation techniques and the related uncertainty work can be separated in three major kinds of activities:

- using nuclear reaction models and experiments as a guide for the nuclear data evaluation: thus uncertainty evaluation is typically a propagation of prior knowledge on nuclear reaction parameters to calculated nuclear data (forward propagation);
- giving greater weight to experimental results (standards), where uncertainty evaluation is mainly a transmission of experimental uncertainties to a function representing the nuclear data;
- using both nuclear reaction models and experimental information with a Bayesian inference procedure where the uncertainty evaluation consists in a mix of prior model knowledge and experimental uncertainties and propagate the results to nuclear data.

One should keep in mind that even though these activities seem quite different, they have common problems to solve: the estimation of proper prior model uncertainties and the importance of taking into account nuclear model deficiencies, experimental uncertainty evaluation that should take into account all sources of uncertainties (electronic counts, background reduction, calibration, detector efficiencies…) and the necessity of finding a proper mathematical description.

A general mathematical description of our problem is to evaluate the following probability density function:

$$p(\vec{\sigma}|\vec{y},\vec{t},\vec{e},U)$$

which is the probability density of the nuclear data (referred to as $\vec{\sigma}$) knowing experimental data (referred to as $\vec{y}$), models (referred to as $\vec{t}$), deficiencies (referred to as $\vec{e}$), see (Leeb et al., 2008; Schnabel, 2015) and prior knowledge (referred to as U). The main differences between the evaluation techniques are focused on the choices of $\vec{t}$, $\vec{e}$ and U as well as the mathematical methods used to estimate this probability density function. Here $\vec{y}$ can be both microscopic experiments (time-of-flight measurements) or integral experiments.

Favouring nuclear reaction models is to choose $\vec{\sigma}\sim\vec{t}$ where $\vec{t}$ is a nuclear reaction model used without direct fitting of experimental information. Sometimes the prior knowledge U is chosen uninformative in this framework. The values taken by $\vec{t}$ can be changed to be compatible with $\vec{y}$, but no explicit adjustment is made.

Favouring experimental results is to choose $\vec{\sigma}\sim\vec{t}$ where $\vec{t}$ is a parameterised function with no physical meaning where the experimental information is taken into account via an adjustment of this function on measurements.

Bayesian inference is a mathematical method that tries to take advantage of both experimental and theoretical knowledge and potential deficiencies. Evaluation of $\vec{\sigma}$ is made through a rigorous mathematical process based on solving the following equation:

$$p(\vec{\sigma}|\vec{y},\vec{t},\vec{e},U) \propto p(\vec{t}|U) \times p(\vec{y}|\vec{t},\vec{e},U)$$

where $\vec{t}$ is a nuclear reaction model whose a priori knowledge is uncorrelated with experiment $\vec{y}$ and deficiencies $\vec{e}$. $p(\vec{y}|\vec{t},\vec{e},U)$ is referred in the community as the likelihood.

Various possibilities are proposed by the evaluator for solving this equation. Deterministic solutions where a multivariate assumption is made for all probability density functions associated with numerical approximations (neglecting second orders) were traditionally used in the evaluation of cross-section in the resonance range for SAMMY code (Larson, 2008), for REFIT (Moxon et al., 2010), and for CONRAD (Archier et al., 2014a), even though these were recently used as well in the higher energy range (Archier et al., 2014b) on a full energy range sodium evaluation. It consists of finding the minimum of a cost function (a generalised chi-square) that contains the a priori information and the theoretical to experimental deviation. For example, if the model $\vec{t}$ contains parameters $\vec{x}$ to be adjusted on experiment $\vec{y}$, the cost function can be the following:

$$\chi^2_{GLS} = \left(\vec{x}-\vec{x}_m\right)^T M_x^{-1}\left(\vec{x}-\vec{x}_m\right) + \left(\vec{y}-\vec{t}\right)^T M_y^{-1}\left(\vec{y}-\vec{t}\right)$$

Where $\vec{x}_m$ is the prior values of the parameters and M_x the associated covariance matrix and M_y the experimental covariance matrix (de Saint Jean et al., 2018) for simple mathematical explanations. $\vec{y}$ can still be both microscopic experiments (time-of-flight measurements) or integral experiments. For the latter, the model $\vec{t}$ is related to additional neutronic calculations with transport code (deterministic codes or Monte Carlo codes) (e.g. Palmiotti et al., 2011; de Saint Jean et al., 2010; Rochman et al., 2019).

Monte Carlo solutions were recently developed allowing a more appropriate framework for this kind of problem (Capote and Smith, 2008; Koning, 2015; de Saint Jean, 2017 for examples and mathematical descriptions). Weighting methods where a priori information is sampled and weighted by the likelihood allowing the evaluation of the expectations and the moments of the posterior distribution were originally proposed (Capote and Smith, 2008). For example, keeping the same notations, one can calculate expectation value for parameters as:

$$< x_i >_{n_k} = \frac{\sum_{k=1}^{n_k} x_{i,k}\mathfrak{L}(\vec{y}_k \mid \vec{x}_k)}{\sum_{k=1}^{n_k} \mathfrak{L}(\vec{y}_k \mid \vec{x}_k)}$$

Where k runs over the sampled prior distribution and $\mathfrak{L}$ is the likelihood value calculated for each realisation k of $\vec{x}$.

The use of weights $\mathfrak{L}_k$ is the major draw-back of this classical BMC method: if the prior value is far from high likelihood values and/or by nature, $\mathfrak{L}_k$ could be dramatically small and the algorithm will have difficulties converging because sampled covered phase space is not favourable. In principle, and idealistically, one should sample with a trial function close to the posterior distribution (Koning, 2015; de Saint Jean et al., 2017). It is worth pointing out that this kind of algorithm will provide an expectation of the first two moments of the posterior probability density function but not the distribution itself. To obtain the final distribution, Markov Chain Monte Carlo based Bayesian inference was proposed (Koning, 2015; de Saint Jean, 2018). This latest method can be considered as a reference calculation as no approximations are made and the only questionable choices are associated

to the ingredient used in the process such as the type of prior distribution (informative, non-informative, Gaussian, log-normal…).

Before going into details in the following chapters, we must emphasise the importance of having proper documentation and descriptions for all ingredients involved in the uncertainty evaluations process: experimental descriptions for all uncertainties, theoretical models used, mathematical methods and the related assumptions (on PDFs) or approximations.

A first issue addressed by the exposed methodologies is the proper treatment of experimental data, especially systematic uncertainties. It is a long-standing problem which faces two major difficulties. First, the lack of experimental description, especially for older experiments, demands additional work for the evaluator to assess the level of these uncertainties (see dedicated chapter in this document). Second, this lack of experimental description raises some additional problems during the adjustment process, including issues such as the proper treatment of systematic uncertainties (Fröhner, 2003; Leeb et al., 2008; de Saint Jean et al., 2009; Habert et al., 2010; Neudecker et al., 2012).

The use of integral experiments for the uncertainty estimation will be discussed, with an emphasis on cross-correlations.

In addition, Monte Carlo solutions give rise to questions on the uncertainty representations, which can be covariance matrices and/or sampled distributions (sampled files). This is not a simple matter, as using covariance matrices is fast but with Gaussian assumptions and using sampled distributions can be considered as reference. The question of the size of the sample is very complex as the result can be very different as a function of the applications and its sensitivity to different nuclear data. For example, one can understand that a relatively low number of files can be used to represent nu-bar distributions and calculate the impact on a k_{eff} of a critical assembly. On the contrary, if the reactor physicist is looking to reaction rates in a special location in the core, one can understand as well that a different number of files should be used for angular distributions or for some inelastic cross-sections in that case.

2.2. Automatic selection and interpretation of EXFOR

Ideally, the full experimental covariance matrix, $\boldsymbol{C}_{\mathbf{exp}}$, including all inter-experiment correlations (normalisation and shared features between experiments) and intra-experiment correlations (target thickness, beam strength, detector calibration, etc.), would be available directly in EXFOR. This is, however, not the case, and an analysis of experimental data results, investigating inter- and intra-experiment uncertainties and correlations as outlined above, is desirable. However, for large scale nuclear data evaluations, such as the TALYS-based Evaluated Nuclear Data Library (TENDL), where a detailed analysis of individual error components of different experiments is not possible, automatic procedures for treating the information available in EXFOR is necessary. For recent versions of TENDL (Koning et al., 2019) this procedure is based on three main components:

- A statistical verification and validation of the EXFOR database (NEA, 2017), where quality flags are assigned to different experiments that are either accepted or rejected. The method is based on a comparison between experimental data and existing nuclear data libraries.
- An assumption of very strong intra-experiment correlations (Koning, 2015). This is obtained by letting each experiment have the same weight when calculating chi-square, independent of the number of experimental data points in this specific

experiment. The procedure does not make full use of the available information in EXFOR on the type of uncertainty (random or systematic).

- Using the entire EXFOR database and TALYS default calculations, to estimate TALYS default uncertainties. This can subsequently be used to create pseudo-experimental data (Koning, 2015; Koning et al., 2019).

The TENDL production is an example where automatic methods are used to interpret the entire EXFOR database to produce a full evaluation for all nuclides using a set of assumptions. Is should be noted, however, that the TENDL evaluation is complemented in some cases by manual selection of experimental information (Koning et al., 2019).

Another attempt for automatic interpretation of EXFOR was proposed by researchers at the University of Uppsala (Helgesson et al., 2017a). Here, the full experimental covariance matrix, $\boldsymbol{C}_{\mathbf{exp}}$, was constructed for specific isotopes. Some of the information for building $\boldsymbol{C}_{\mathbf{exp}}$ is available in EXFOR and in the XC4 format; however, the information is not complete, and a set of rules was constructed to automatically interpret EXFOR. The overarching idea for the rules is to penalise experiments with unrealistic or badly reported uncertainty. The work showed that, given a set of assumptions dictated in the rules, $\boldsymbol{C}_{\mathbf{exp}}$ could be constructed, which could subsequently be used in the nuclear data evaluation. The work also contained sensitivity studies of the assumptions, and it was found that nuclear data applications were only weakly affected by many of the assumptions. In contrast, some of the assumptions had substantial effects on the investigated applications. The simplified version of the rule-based approach for building $\boldsymbol{C}_{\mathbf{exp}}$ has subsequently been integrated into a prototype pipeline for ND-evaluation and production of an example evaluation of ^{56}Fe (Schnabel et al., 2021). A set of less elaborate rules has been applied when interpreting the EXFOR database (Schnabel, 2015). In other approaches (Helgesson et al., 2017a), a similar but slightly different method to what is applied for TENDL (NEA, 2017) for rejecting experiments was also applied. It was found that the resulting nuclear data evaluation and its impact on the application were particularly sensitive to the choice of rejection criteria.

It is clear from the above examples that there is no consistent way to automatically interpret the EXFOR database. The choices and assumptions made when constructing the different evaluation system will significantly affect the final evaluation and the associated covariance file. Therefore, the proposed subgroup on developing an automatically readable, comprehensive and curated experimental reaction database could significantly streamline the automatic interpretation of experimental data and provide quality assurance.

The selection criteria described in this section are based on the deviation between theory/evaluations and experiments. Ideally, the selection criteria should also/instead be based on intercomparison between experimental datasets (NEA, 2017; Helgesson et al., 2017a). In addition, the proposed and used methods are binary (accept/reject) and, ideally, a selection mechanism should be used that can gradually reweight the importance of different experiments, depending on their agreement with other experiments. This is addressed in the subsequent section.

2.3. Handling of discrepant data sets

When retrieving data from EXFOR, the data sets are often discrepant. Even after a detailed analysis of experimental data as outlined in 3a, data might still be inconsistent due to hidden error components/Unrecognised Sources of Uncertainties (USU) (Capote et al., 2020). The USU are the unexplainable differences between datasets after a careful analysis of the experimental data. For more automatic treatment of experimental data, as reported in the

previous section, the existence of USU can also be due to failure to include all uncertainty components reported in the experiment in the EXFOR database in a consistent way.

USU lead to an erroneous experimental covariance matrix, which in turn can severely affect the evaluated covariance (Helgesson et al., 2018) or as stated by Capote et al. (2020) "if one neglects adding obviously necessary contributions from USU to uncertainties of input data the evaluated uncertainties will be underestimated, in turn, adversely impacting application calculations". There is therefore a strong motivation for treating USU.

The sources of USU can be "unrecognised uncertainty across many data sets due to using the same methods", "Missing cross-correlations between experimental data" or "Missing uncertainty sources for single experimental data sets" (Neudecker et al., 2018). Treating USU means estimating the magnitude of the USU and possible correlations between the USU, and from this constructing an extra covariance matrix, $\boldsymbol{C}_{\mathbf{usu}}$, which is added to the $\boldsymbol{C}_{\mathbf{exp}}$ from the recognised sources of uncertainties.

There are several ways to treat USU as outlined by Capote et al. (2020). SG-44 has particularly highlighted the maximum/marginal likelihood estimation/optimisation (MLO) (Schnabel, 2017; Sjöstrand and Schnabel, 2019; Capote et al., 2020) and the so-called "Physical Uncertainty Bounds" (PUBs) (Vaughan and Preston, 2015; Neudecker et al., 2018; Capote et al., 2020).

The PUB method effectively gives clues to understand discrepant data sets. The MLO method finds the most likely $\boldsymbol{C}_{\mathbf{usu}}$ given the experimental data. The MLO is based on an inter- and intra-comparison between experiments. This is in line with what was recommended in recent reviews (NEA, 2017; Helgesson et al., 2017a) and outlined in the previous section. Large uncertainties are added to data points that deviate from the majority of experiments. This results in a "soft" rejection of outliers. This is also perceived as a strength of the method since it is not binary and other methods for rejection of data can be abandoned. Normally, some structure of the $\boldsymbol{C}_{\mathbf{usu}}$ must be assumed (Capote et al., 2020). In the example evaluation generated by Schnabel, systematic intra-experiment uncertainties were assumed but no inter-experiment uncertainties. It is typically argued that inter-experiment uncertainties should be assigned based on the type of experiment. The method can be employed both when the nuclear data PDF is reported in a covariance file as well as if the nuclear data PDF is reported using random files. It can both be used for differential data as well as integral data (Siefman, 2019; Sjöstrand and Schnabel, 2019). The method can also incorporate prior information on the magnitude of the USU. The data-driven nature of the MLO makes it particularly suited to be employed for automatic ND evaluations and it is hence also integrated into a prototype pipeline for nuclear data evaluation.

There is a wealth of literature in the scientific community outside the nuclear data (ND) community on handling discrepant data and outliers. SG-44 has explored a few options and can conclude that there is a need to treat discrepant data and USU and that the choice of assumptions will affect the resulting evaluated covariance. Hence, the recommendation outlined by Capote et al. (2020) should be followed when constructing methods for treating USU: "As few arbitrary assumptions as possible should be involved" and "well-defined, transparent mathematical algorithms should be employed" so that the procedures involved "deliver results that could be replicated by a future evaluator".

3. Use of integral experimental data

Integral experiments are essential for the validation of nuclear data libraries for practical applications. Currently, in the international nuclear data community, critical integral experiments in particular are heavily used in the validation process. There is almost universal agreement and desire for more kinds of integral (and semi-integral) experiments to be included in the validation process. WPEC Subgroup 47 is working on bringing forth a wider use of shielding and transmission experiments. There is also increased support for incorporating validation data from commercial power reactors. While both of these directions are supported by this subgroup, fully adopting these new sources of validation data will take time.

There is a wide spectrum of views in this subgroup and in the international nuclear data community on how validation data, in particular critical integral experiments, influence the final product of the evaluated nuclear data libraries. Some argue that the validation data have no influence on the evaluated nuclear data and only serve as a final check. Others are of the view that nuclear data evaluators are influenced by the results of the validation checks before the final evaluation is submitted to the nuclear data library. There are also many views on how the evaluators use this knowledge. Some argue that it is acceptable for nuclear data evaluators to check the performance of trial versions of their evaluation on some subset of the validation suite, while others see rejection of trial versions of isotope evaluation due to poor performance in integral tests as assimilation of information from integral experiments. There is also a view that integral data could or should be used right away by the evaluators to produce the best possible evaluation, which simultaneously satisfies both differential and integral measurements (Bauge and Rochman, 2018; de Saint Jean et al., 2015; de Saint Jean et al., 2018; Rochman et al., 2017) for the use of integral experiment for covariance estimation. Further, the boundary between differential measurements, thick-sample transmission measurements, semi-integral measurements, single-material shielding measurements and integral measurements is not universally defined. This is true with respect to the mean values of the evaluation.

On the other hand, there is almost universal agreement that the evaluation of nuclear data uncertainty in the current nuclear data libraries is based only on the uncertainty in the differential data. This was reflected by the statement released with the ENDF/B-VIII.0 library:

- The covariance data in the ENDF evaluations represent uncertainties and correlations in differential data.
- The use of this covariance to calculate uncertainties for integral quantities such as k_{eff} will usually result in an overestimation of the uncertainty. That said, comparisons with integral data are essential during the evaluation process and users should not be surprised if the *mean value* nuclear data allow for the accurate prediction of k_{eff}, even if the covariances do not reflect this consideration.
- The recommended methodology to overcome this problem is to adjust the covariance to add information from a set of integral data that represents the physics of the system for which the adjusted covariance will be used.
- More information on this topic is available at: www.oecd-nea.org/download/wpec/sg33/.

- The US DOE Cross Section Evaluation Working Group (CSEWG) is currently studying the best covariance representation for future releases.

This subgroup could not come to a consensus on a recommendation on the proper use of integral experiments in nuclear data evaluations, although the debates were robust. A new subgroup was opened under the auspices of the WPEC, Subgroup 49, which will, as part of its mandate, explore this issue. One of the current proposals is to have several parallel nuclear data libraries with a different degree of assimilation of information from integral experiments. For example, there could be a library which is solely based, for mean values and covariances, on differential measurements, and a library which is systematically adjusted for a wide (although specified) range of applications based on the coverage of integral experiments.

As can be seen from the ENDF-covariance disclaimer above, currently the evaluated nuclear data libraries are somewhere between the two extremes considered by Subgroup 49. For some members of this subgroup, the practice of the "comparisons to integral data" during the evaluation process brings up the question of consistency between the mean values and their associated covariance. While some argue that the covariances can be reconciled with the mean values by the systematic assimilation of integral data, others say that the prior covariance going into this assimilation is already inconsistent.

In contrast, there was near universal agreement in this subgroup that any use of integral experiments in the evaluation process is rarely documented in current practice. While this subgroup cannot recommend whether, or how, integral experiments should be used in the evaluation process, it does strongly recommend that any use of integral experiments be documented in the evaluation publication. It recommends journal editors to demand from authors documentation on the use of integral experiments with the following details.

There should be documentation of which integral experiments were considered by the evaluator prior to the submission to the library for testing, regardless of how the integral information was used. Examples include, but are not limited to:

- integral experiments used in the validation of a nuclear data library which have motivated a new or re- evaluation of an isotope, reaction or energy region;
- integral experiments used by the evaluator for testing and/or rejecting several trial versions of evaluations;
- integral experiments used by the evaluator for systematic or non-systematic adjustment of evaluations.

Furthermore, integral experiment evaluation revisions should be documented along with input-deck sources, computer codes and any corrections or adjustments to integral evaluations.

It is also important to clearly document how the integral experiments were used, even if they were used only to reject poorly performing trial evaluations. Ideally, if used by individual evaluators, integral experiments should be used in a systematic procedure that is reproducible. Reproducibility will be one of the major themes of Subgroup 49. However, any use needs to be documented and transparent to allow for a full understanding and appreciation of what was done to produce the nuclear data evaluation.

3.1. Other probability distributions for uncertainties

A long-standing problem in nuclear data covariance is the assumption of Gaussian distribution for inherently positive quantities. Currently, the Gaussian distribution is used

exclusively to communicate nuclear data uncertainty information in the ENDF-6 format. While propagation of uncertainty through linear (or linearised systems) under the Gaussian assumption is mathematically convenient, problems arise for random sampling. Nuclear data quantities with large uncertainties, for example 40%, have a significant probability of being sampled as a negative value. This, of course, does not make physical sense and is difficult to handle for simulation codes.

The new Generalised Nuclear Data Structure (GNDS) format (NEA, 2020b), the subject of the WPEC Expert Group on GNDS, allows the evaluator to define different probability distributions. However, a universal prescription for the fix-up for covariance evaluations already described as Gaussian in the library but with a large uncertainty has not been established and different practitioners correct for this in different ways. Below, several common fixes are summarised and their advantages and disadvantages described based on the report of Sébastien Lahaye at the 2017 Covariance Workshop, entitled "Choice of positive distribution law for nuclear data" (Lahaye, 2018).

3.1.1 Truncated Gaussian

A truncated Gaussian is achieved simply by sampling the original Gaussian distribution as defined by the mean and variance (covariance) but rejecting any negative samples. This is perhaps the most common fix; however, it must be recognised that the mean and variance of the resulting distribution are no longer the same as that of the original. For a decreasing ratio of the mean to the standard deviation of the original distribution, the mean of the new distribution will be larger and the variance smaller than that of the original. On the other hand, the behaviour of the positive tail of the distribution will be most similar to that of the original from all of the other fix-up alternatives listed below.

3.1.2 Fitted Gaussian

A fitted Gaussian is also truncated at zero by rejecting samples below zero. However, for this distribution, sampling is done not from a Gaussian with the original mean and variance as listed in the nuclear data file but from a lower mean and larger variance, such that the resulting distribution, which lacks values below zero, will have a mean and variance of that listed in the evaluated file. This approach has the advantage that the first two moments of the distribution are preserved. However, the behaviour of the distribution can be notably different at the peak for low values of the ratio of mean-to-standard deviation.

3.1.3 Log-normal

The log-normal distribution follows from the maximum-entropy principle for inherently positive quantities. The log-normal distribution can be established with the same mean and variance as the original and will guarantee only positive samples. However, the behaviour of the peak and tail of the distribution will be most different for low mean-to-standard deviation ratios from the other options listed above.

The most pressing issue is that when such cases occur in evaluated nuclear data libraries, it is unclear which distribution was the intent of the evaluator. Therefore, this subgroup recommends that the evaluators make sure that sampling from the evaluated covariance data does not result in negative samples too often. While we hesitate to provide a precise prescription for too often, we simply encourage the evaluators to define a fix-up prescription for the users on how negative samples should be handled.

3.1.4 Random files

One alternative to providing covariance data and defining a distribution is to provide many realisations/samples of evaluated files. This is an attractive alternative, especially when the parameter uncertainties are so large that propagation of uncertainty is not linear. The number of random files used to provide sufficient coverage has been about 1 000. However, when cross-isotope correlations are available, random realisations of entire libraries rather than only single isotope evaluations should be provided.

3.2. Non-Gaussian distributions within GNDS

Covariance matrices generally imply that uncertainties are Gaussian-distributed. This assumption is problematic for nuclear data, however, since quantities like cross-sections, average multiplicities and probability distributions for outgoing energy and angle are constrained to be non-negative. A normal distribution can easily produce negative samples, especially if the mean is small and the variance large.

Users of nuclear data covariances have come up with various strategies to deal with this problem. Some simple strategies include:

- truncating realisations to remove unphysical negative values; and
- generating many random realisations and then keeping only those samples that avoid unphysical negative values.

These approaches are not ideal, however. They may require drawing very many samples (especially for large covariance matrices when cross-terms between many different reactions are present), and they introduce a bias in sampling.

Another frequently-discussed option is to move away from the assumption of Gaussian-distributed covariances in favour of intrinsically positive probability distributions such as log-normal. Log-normal pdfs have the advantage of ensuring that all samples remain positive, but they require that some basic assumptions about covariances be revisited. For example, it has been demonstrated that log-normal covariance matrices can be used to sample resonance parameters even when the relative uncertainties grow large, but also that samples converge very slowly if especially large relative uncertainties are present (Žerovnik et al., 2015).

Version 1.9 of GNDS format (NEA, 2020b) supports storing log-normal uncertainties and covariances, although the option has yet to be exercised in current evaluations.

3.3. Validation of nuclear data uncertainties via "Physical Uncertainty Bounds"

The previous sections deal with how evaluated uncertainties can be propagated to integral experiments to yield uncertainties on application calculations stemming from nuclear data. However, these application calculation-related uncertainties can only be realistic if the input nuclear data covariances are reliable as well.

Of course, a covariance matrix is not a physical quantity in itself that can be right or wrong. However, it can definitely be realistic, under- or over-estimated given the information that was used for the evaluation. If, for instance, three experimental data sets of an uncertainty in the range of 1–5% covering the whole energy range are used as the sole input for an evaluation, an evaluated uncertainty below 0.01% or above 100% would be clearly identified as unrealistic.

Very few methods are used in the field of nuclear data evaluation to routinely assess the reliability of uncertainties given in our libraries. They are usually based on expert judgement. One obvious example to validate nuclear data uncertainties is the following: If a quantity, A, is measured in most cases relative to a specific Neutron Data Standards observable (Carlson et al., 2018), B, then the evaluated uncertainties of A are expected to be larger than the uncertainties of B.

However, more formalised mathematical procedures exist to validate nuclear data uncertainties based on the information actually used for the evaluation. One method that can be used to this end is the "Physical Uncertainty Bounds" (PUBs) method. Vaughan and Preston (Vaughan and Preston, 2015) developed this method to yield (a) upper bounds on the uncertainties of integrated systems/observables, and (b) functional forms of these observables obeying these bounds and basic physics considerations. The original method was applied to many areas in physics including, but not limited to, plasma fusion reactions, material damage or strength. Part (a) of this method was modified such that it allows to validate nuclear data uncertainties obtained primarily on a statistical analysis of experimental data (Neudecker et al., 2020; Capote et al., 2020). The main steps in applying this procedure are summarised below. Also, an example is shown by applying it to validating ENDF/B-VIII.0 versus ENDF/B-VII.1 average ^{239}Pu(n,f) prompt neutron multiplicities, ^{239}Pu $<\nu_p>$, in the incident-neutron energy range from 0.1 to 20 MeV.

The PUBs method, as applied here to validate evaluated uncertainties obtained from a statistical analysis of experimental data, comprises the following steps:

- The independent physics sub-processes pertinent to each specific class of measurement of this observable are identified. Care must be taken to identify truly independent physics sub-processes. If correlations exist between them (e.g. between multiple scattering and attenuation correction procedures that both rely on the same neutron transport codes and underlying data), these sub-processes should be grouped into one.
- Sensitivity studies should be performed to investigate which of these sub-processes yield non-negligible uncertainty-contributions to the total bounds. If even an extreme variation in one physics sub-process has an insignificant impact on the final observable, the dimensionality of the problem can be reduced.
- Minimal realistic and conservative bounds are estimated for each sub-process as related to the observable of interest. These bounds are based on often multiple, reliable experimental data sets, numerical data from first principles, or fundamental theory.
- A functional form relating the observable with each sub-process is quantified, guided by physics constrains, experimental data and numerical data. These functional forms can inform the shape of correlation matrices related to the bounds of each sub-process. Together, these uncertainties and correlation shapes yield minimal realistic and conservative covariances for each sub-process.
- The independent minimal realistic and conservative covariance for each sub-process is summed up in quadrature to two total covariances.

It is key for this analysis that the information used for the validation of uncertainties be as similar as possible to that used for the evaluation. If a certain set of experimental data was used for an evaluation, the same set should be used, if known, for the validation.

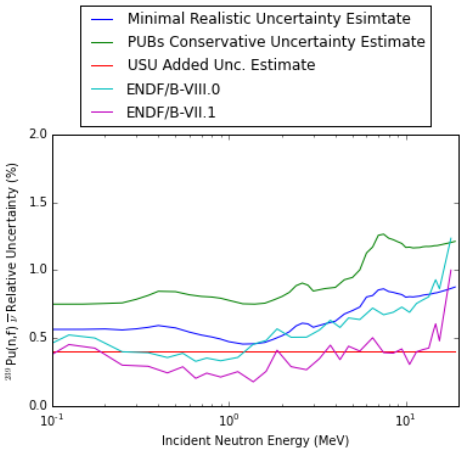
Reliable evaluated variances are expected to be within the uncertainty range spanned by minimal realistic or conservative bounds. If the evaluated uncertainties are smaller than the

minimal realistic bounds, they are likely underestimated. If they are larger than the conservative bounds, they are likely overestimated.

In the following example, PUBs are applied to investigate whether the ^{239}Pu $<\nu_p>$ uncertainties from the US nuclear data libraries ENDF/B-VII.1 (Chadwick et al., 2011) are more or less realistic than those from ENDF/B-VIII.0 (Brown et al., 2018). This particular example was chosen because the associated uncertainties change distinctly from ENDF/B-VII.1 to ENDF/B-VIII.0, as can be seen from Figure 3.1. To be more specific, ENDF/B-VIII.0 uncertainties are systematically larger for all incident-neutron energies compared to their ENDF/B-VII.1 counterparts. This distinct change in the evaluated ^{239}Pu $<\nu_p>$ uncertainties has important implications for application calculations. Uncertainties in Jezebel k_{eff} due to those of ^{239}Pu $<\nu_p>$ nearly tripled from ENDF/B-VII.1 to ENDF/B-VIII.0 from 81 to 240 pcm (Chadwick et al., 2018; NEA, 2019b). While 81 pcm are less than one-third of a dollar (270 pcm) in Pu systems, ENDF/B-VIII.0 indicates that we do not understand this observable accurately enough to predict k_{eff} of Jezebel within more than 85% of a dollar. This result might imply that we need to undertake new evaluation and, possibly, experimental efforts to decrease the ^{239}Pu $<\nu_p>$ uncertainties.

One major problem of this analysis was that it was not know which data sets were used for the evaluations. Hence, all data sets that are available for the ^{239}Pu $<\nu_p>$ in EXFOR were analysed. A total of 26 out of these 30 experiments were used as input for the PUBs analysis. These were the experimental data either partially in the energy range of interest or correlated to one of those measurements. Only a subset of these 26 data sets provided information on the relevant physics sub-processes. In addition, information was taken from other sources (Capote et al., 2020; Boldeman and Fréhaut, 1980; Lovell et al., 2020) to estimate bounds on the individual physics sub-processes. Thirteen independent sub-processes were identified that contribute to uncertainties on ^{239}Pu $<\nu_p>$ measurements: background in the neutron detector, ^{252}Cf standard, corrections for false fission events, correction for anisotropic fission fragment emission, multiple scattering corrections, counting statistics, energy resolution, differences in the prompt fission neutron spectra entering the neutron detector efficiency calculations, delayed gammas causing a background, foil thickness, dead-time, impurities in the sample and displacement of fission sample. Of these 13 sub-processes, the last 5 contribute only little to the uncertainty budget with uncertainties in the range of 0.01–0.15%. The other uncertainty sources must be accounted for in an uncertainty estimate of these types of measurements. Nonetheless, minimal realistic and conservative PUBs bounds were estimated for each of those 13 sub-processes along with shapes of associated correlation matrices. The total minimal realistic and conservative PUBs uncertainties in Figure 3.1. are then obtained by summing the respective bounds of all independent sub-processes in quadrature. This result is based only on experimental information as, in the particular energy range chosen, the ENDF/B-VII.1 and ENDF/B-VIII.0 evaluations are mainly based on experimental data.

Figure 3.1. The minimal realistic and conservative ^{239}Pu <νp> PUB uncertainties compared to their ENDF/B-VII.1 and ENDF/B-VIII.0 counterparts



Source: LANL, 2021.

The PUBs results in Figure 3.1. indicate that ENDF/B-VII.1 ^{239}Pu <ν_p> uncertainties are underestimated below 10 MeV. The increase of evaluated uncertainties for ENDF/B-VIII.0 was indeed justified. However, even these seem unrealistically low below 1 MeV and low from 1–10 MeV compared to minimal realistic and conservative PUBs uncertainties.

3.4. Available integral experiment correlation data

Two main sources of integral experimental benchmarks to validate nuclear data are the International Criticality Safety Benchmark Evaluation Project (ICSBEP) (NEA, 2019a) and International Reactor Physics Evaluation (IRPhE) Project (NEA, 2019c). Correlation data between the benchmark model uncertainties is limited; 94 cases, or ~2% of the nearly 5 000 ICSBEP cases, have numerical correlation (Ivanova et al., 2003) data that are retrievable via the **D**atabase for **IC**SB**E**P (DICE) (Hill et al., 2013). In addition to quantitative correlation data, many evaluations provide a binary qualitative indicator of other evaluations that are correlated, indicating whether the same fuel or tank has been used. In DICE these qualitative correlations have assumed the property of transitivity.

Figure 3.2. Example of quantitative (left) and qualitative (right) correlation coefficients in the NEA DICE database

	HMF 008 001	HMF 011 001	HMF 018 001	HMF 020 001	HMF 031 001	HMF 055 001	HMF 060 001	HMF 061 001	HMF 067 001	HMF 067 002	HMF 070 001	HMF 070 002	HMF 070 003	HMF 075 001	HMI 001 001
HMF008-001	1000	210													
HMF011-001	210	1000													
HMF018-001			1000	460	320										
HMF020-001			460	1000	460										
HMF031-001			320	460	1000										
HMF055-001						1000	300	250	290	290	260	250	270	210	210
HMF060-001						300	1000	510	880	880	840	840	850	430	680
HMF061-001						250	510	1000	500	500	440	430	450	870	370
HMF067-001						290	880	500	1000	960	930	940	940	420	770
HMF067-002						290	880	500	960	1000	940	940	940	420	780
HMF070-001						260	840	440	930	940	1000	940	930	370	780
HMF070-002						250	840	430	940	940	940	1000	940	360	800
HMF070-003						270	850	450	940	940	930	940	1000	380	790
HMF075-001						210	430	870	420	420	370	360	380	1000	310
HMI001-001						210	680	370	770	780	780	800	790	310	1000
HMM012-001						270	540	760	520	520	470	460	480	810	380

	HCF 001	HCF 002	HCF 004	HCI 003	HCI 004	HCI 005	HCM 003	HCM 004	HCT 003	HCT 004	HCT 005	HCT 006	HCT 007	HCT 008
HCF001	(+)	+	+											
HCF002	+	(+)	+											
HCF004	+	+	(+)											
HCI003				(+)										
HCI004					(+)									
HCI005						(+)								
HCM003							(+)	+						
HCM004							+	(+)						
HCT003									(+)	+	+	+	+	+
HCT004									+	(+)	+	+	+	+
HCT005									+	+	(+)	+	+	+
HCT006									+	+	+	(+)	+	+
HCT007									+	+	+	+	(+)	+
HCT008									+	+	+	+	+	(+)
HCT009							(+)	+						
HCT011														

Source: NEA, 2021.

In the reactor physics handbook, correlation coefficient data have been provided for the recent KRITZ-LWR-RESR-004 experimental benchmark on temperature reactivity

coefficients. In reactivity effect experiments, determining a correlation between two states is a prerequisite to accurately assessing the experimental benchmark reactivity.

The final report of WPEC Subgroup 33 provided practical guidance and illustrative examples of determining and using experimental benchmark correlations in an adjustment process (NEA, 2013).

3.4.1. Experimental error matrix

The experimental error values of an integral parameter are usually given by the experimenters with the error components. However, the correlations between multiple integral parameters are rarely found in the experiment report; therefore, these correlation factors must be estimated from the available experimental information. The error component correlation method adopts the following three steps.

Stage 1 - Classification of error components to either common or independent

First, all related components of the experimental errors for "Data A" and "Data B" with quantitative values reported are listed, and each individual component identified either as a "common error (i.e. the correlation factor is 1.0) between Data A and B", or an "independent error (i.e. the correlation factor is 0.0)"[1]. If an error component is judged as a mixture of common and independent errors, that is, the correlation factor is not considered as either 1.0 or 0.0, the error component must be divided into more detailed subcomponents until the error component becomes either a common or independent error. This classification requirement is difficult for the experimenters who evaluate the error components in their report, but today this kind of rigour is essential to retain full value of these experimental quantities. Recent experimental databases like the NEA ICSBEP and IRPhE handbooks now include such detailed experimental error evaluation due to the continuous efforts of the authors and reviewers.

Stage 2 - Summation of common and independent errors

Next, the common and independent errors, respectively, are summed by the statistical method, that is, the "square, sum and root" means to obtain each standard deviation, σ_{Total}, the diagonal term of matrix. The statistical treatment is justified by the assumption that all error components have already been divided until there are no correlations between any error items in the measurement of an integral parameter. The total errors of Data A and B, that is, the diagonal term of error matrix, Ve, are the sum of squared common and independent errors.

Stage 3 - Evaluation of correlation factor

Finally, the correlation factor, non-diagonal term, of Data A and B is derived as the ratio of common errors to the total errors as the equation below. The "Stage 1 to 3" procedures must be repeated for all matrix elements to generate a full experimental error matrix as the input of adjustment exercise. It should be noted that the correlation factors between several

1. The words "common" and "independent" adopted here are usually referred as "systematic" and "statistical", respectively, in many experimental reporting literatures. However, the use of the former labels more clearly expresses the intention of this classification to evaluate their correlation factor for a specific pair of data in a large matrix than the latter labels.

sodium void reactivity measurements would be changed depending on the combination of void steps, even in the same experimental core, so that

$$\rho_{A,B} = \frac{\sum_i \sigma_{Common,A,i} \times \sigma_{Common,B,i}}{\sigma_{Total,A} \times \sigma_{Total,B}},$$

where, the suffix *i* represents common error components between Data A and Data B. Examples can be found in previous WPEC Subgroup 33 reports (NEA, 2013).

3.4.2. Template ICSBEP uncertainties and semi-automatic generation of within evaluation covariance

The impact of integral experimental benchmark correlations on applications and nuclear data adjustment has not been extensively examined. Efforts in the area of integral experiment correlations have been relatively neglected when contrasted against the advances in nuclear data covariance between 2005 and 2020.

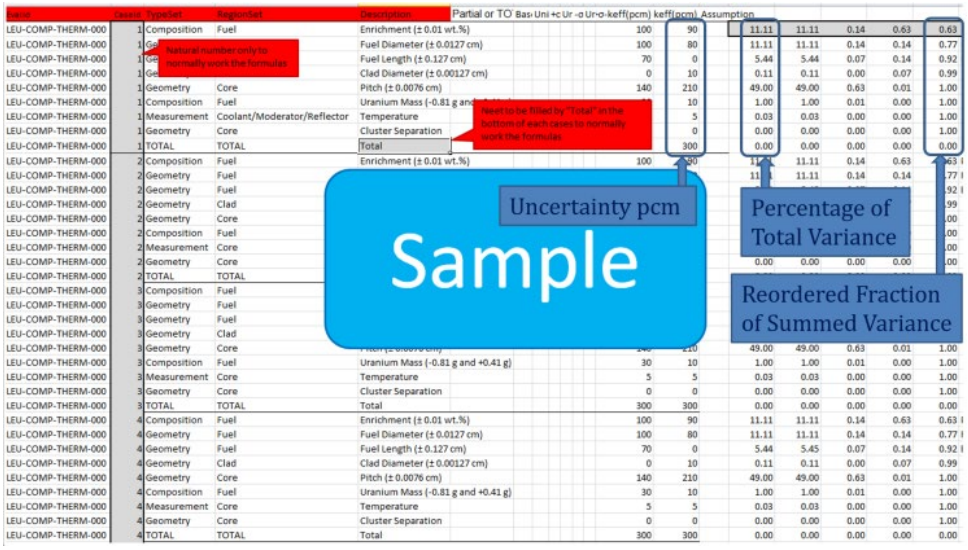
With 4 800 ICSBEP cases lacking experimental correlations, a significant effort is required to fill in the remaining experimental correlation data. Unfortunately, in-depth knowledge of the experiment is required to determine if each component is shared or independent between cases; this information has often not been recorded in Section 2 of the ICSBEP evaluations and so judgements based on the recorded information can only assess the probability of being shard or not.

To facilitate the assignment of experimental correlations a template was created (Jeong et al., 2017) to determine the correlation coefficient between cases within an evaluation. The tool is similar to a tool made for nuclear data uncertainties by Zerkin (Zerkin, 2012). To populate the template, the first step is to extract all of the ICSBEP Evaluation Sections 2 and 3 uncertainties from the benchmark evaluations. Evaluations often provide detailed uncertainty estimates only for a subset of the cases, so ensuring correct matching between surrogate and individual cases is necessary. Furthermore, many evaluations do not have a simple summary table, so extracting the correct uncertainty values is non-trivial, requiring familiarity with ICSBEP evaluations and formats.

With the list of all uncertainty components, the excel template can compute the sum of individual uncertainties in quadrature and compare it to the total uncertainty, flagging any large disagreements for examination.

Next, the percentage variance that each uncertainty term contributes is computed and the most influential terms are identified. The template currently applies a criterion to identify the top contributing terms corresponding to at least 90% of the total variance; specifically, 90% of the sum of the individual variances (rather than comparing to the previous section). Users then can decide whether these top contributors are shared between cases, allowing for the total shared uncertainty between the two cases to be computed. Currently the template allows only for covariances between cases within evaluations, although it is recognised that in the future, functionality for inter-evaluation case level correlations will be needed.

Figure 3.3. Template for ICSBEP uncertainties



Source: NEA, 2021.

Figure 3.4. Template for ICSBEP uncertainties (II)

1	**Comments:**							**Correlations(Betwee**	All Cases									
2								Pitch (± 0.0076 cm)	0.99									
3	**Average C/E**							Fuel Diameter (± 0.0	0.99									
4	0.99911883							Enrichment (± 0.01 w	0.99									
5								Fuel Length (± 0.127	0.99									
6	**Weighted Average C/E**							Cluster Separation	0.2									
7								Free	0									
8								Free	0									
9	**GLS Average C/E**							Free	0									
10	0.99437309							Free	0									
13								**Uncertainties(pcm)**	**Case1**	**Case2**	**Case3**	**Case4**	**Case5**	**Case6**	**Case7**	**Case8**		
14								Pitch (± 0.0076 cm)	210	210	210	210	210	210	210	210		
15								Fuel Diameter (± 0.0	100	100	100	100	100	100	100	100		
16								Enrichment (± 0.01 w	100	100	100	100	100	100	100	100		
17								Fuel Length (± 0.127	70	70	70	70	70	70	70	70		
18								Cluster Separation	0	40	50	50	50	50	90	40		
19								Free	0	0	0	0	0	0	0	0		
20								Free	0	0	0	0	0	0	0	0		
21								Free	0	0	0	0	0	0	0	0		
22								Free	0	0	0	0	0	0	0	0		
23								Free	0	0	0	0	0	0	0	0		
24								Total	264.6	267.6	269.3	269.3	269.3	269.3	279.5	267.6		
25								**Uncertainty Matrix (pcm sigma)**										
26	**Uncertainties(p**	Pitch	Fuel Diameter	Enrichment	Fuel Length	Cluster Separation	Total		LCT001-1	LCT001-2	LCT001-3	LCT001-4	LCT001-5	LCT001-6	LCT001-7	LCT001-8		
27	Case1	210	100	100	70	0	264.6	LCT001-1	264.6	261.4	261.4	261.4	261.4	261.4	261.4	261.4		
28	Case2	210	100	100	70	40	267.6	LCT001-2	261.4	267.6	262.1	262.1	262.1	262.1	262.7	262.0		
29	Case3	210	100	100	70	50	269.3	LCT001-3	261.4	262.1	269.3	262.3	262.3	262.3	263.1	262.1		
30	Case4	210	100	100	70	50	269.3	LCT001-4	261.4	262.1	262.3	269.3	262.3	262.3	263.1	262.1		
31	Case5	210	100	100	70	50	269.3	LCT001-5	261.4	262.1	262.3	262.3	269.3	262.3	263.1	262.1		
32	Case6	210	100	100	70	50	269.3	LCT001-6	261.4	262.1	262.3	262.3	262.3	269.3	263.1	262.1		
33	Case7	210	100	100	70	90	279.5	LCT001-7	261.4	262.7	263.1	263.1	263.1	263.1	279.5	262.7		
34	Case8	210	100	100	70	40	267.6	LCT001-8	261.4	262.0	262.1	262.1	262.1	262.1	262.7	267.6		
36									**X'**									
37									**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**		
38									**LCT001-1**	**LCT001-2**	**LCT001-3**	**LCT001-4**	**LCT001-5**	**LCT001-6**	**LCT001-7**	**LCT001-8**		
39								**LCT001-1**	1.000	0.965	0.959	0.959	0.959	0.959	0.924	0.965		
40								**LCT001-2**	0.965	1.000	0.953	0.953	0.953	0.953	0.923	0.958		
41								**LCT001-3**	0.959	0.953	1.000	0.949	0.949	0.949	0.919	0.953		
42								**LCT001-4**	0.959	0.953	0.949	1.000	0.949	0.949	0.919	0.953		
43								**LCT001-5**	0.959	0.953	0.949	0.949	1.000	0.949	0.919	0.953		
44								**LCT001-6**	0.959	0.953	0.949	0.949	0.949	1.000	0.919	0.953		
45								**LCT001-7**	0.924	0.923	0.919	0.919	0.919	0.919	1.000	0.923		
46								**LCT001-8**	0.965	0.958	0.953	0.953	0.953	0.953	0.923	1.000		

Source: NEA, 2021.

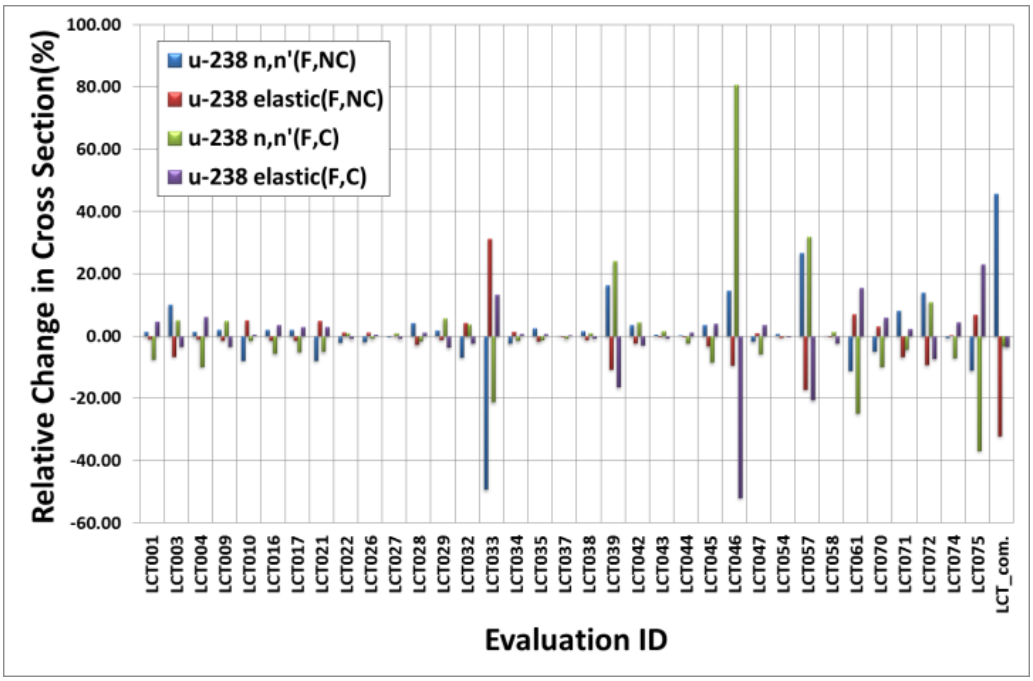
3.4.3. Impact of integral covariance

Using the TSURFER routine in SCALE-6.2 (Wieselquist et al., 2020) the impact of an experimental covariance matrix was analysed. The tests were simplistic studies running the code with and without experimental correlations and seeing the impact on ^{238}U elastic and inelastic fast energy cross-section adjustments.

The input data for all the runs consisted of sensitivity profiles computed using the 238-group ENDF/B-VII.0 library distributed with SCALE and the 44-energy group nuclear data covariance library. The C/E values were done with continuous energy KENO using the ENDF/B-VII.0 library as these calculated results had a lower uncertainty from modelling approximations. Results presented in the following sections comprise the standard available output from a TSURFER calculation. Adjustments were performed within an experimental series, i.e. running surfer for just LCT001 with and without correlations between cases, and for all LCT cases with and without correlation.

Figure 3.5 shows the results for ^{238}U inelastic and elastic cross-section in the fast energy region. It is informative to compare the proposed LCT adjustment with and without correlations. The result is significantly different, as the uncorrelated adjustments are a large 40% increase for inelastic and a 40% decrease for elastic, while the correlated adjustments are a few percent and in the same direction. Clearly correlation coefficients have a large impact on any proposed adjustments, and need to be fully considered for any adjustment or optimisation method to provide feedback.

Figure 3.5. Impact of integral experiment correlations on TSURFER proposed adjustments to ^{238}U capture and inelastic cross-sections



Source: NEA, 2021.

4. Cross-correlations

It is intuitive for the cross-section of all of the isotopes of an element to be correlated through the evaluation process as it is often the case that experimental measurements on natural samples are being evaluated. Therefore, the uncertainty information of the cross-section of the individual isotopes is all tied together through the cross-isotope correlations established in the analysis methodology. The joint covariance matrix is often calculated by using the generalised linear least-squares (GLLS) updating technique implemented in SAMMY:

$$M' = (M^{-1} + G^t V^{-1} G)^{-1},$$

where M' is the posterior covariance matrix, M is the prior covariance matrix, V is the covariance matrix for the experimental data being analysed, and G is the sensitivity matrix of the observable measured to the quantity being evaluated. The prior covariance matrix, M, may be block-diagonal with respect to the individual isotopes of the natural element. However, the posterior covariance matrix, M', is a full matrix.

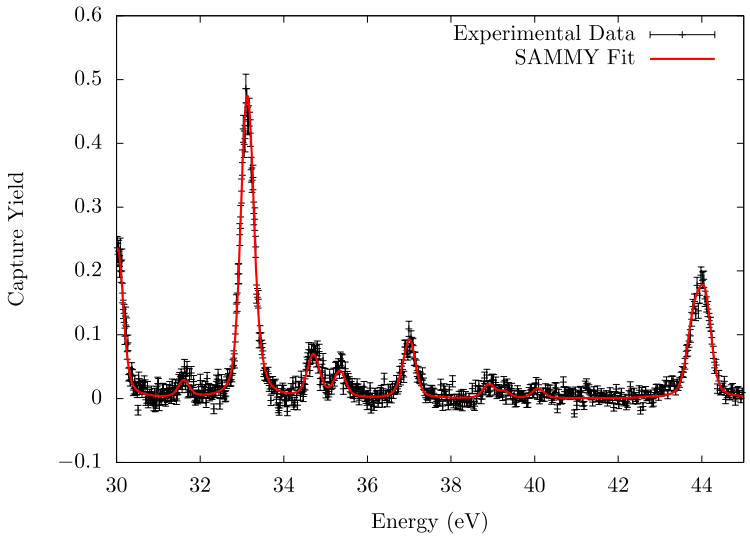
A covariance matrix, including the cross-isotope covariances, is necessary to accurately reflect the state of knowledge of the cross-section simultaneously, both for the individual isotopes and for their combination in the calculation of the cross-section of the natural element. Only through reporting cross-isotope covariances can the evaluator accurately reflect the fact that the cross-section for the natural element (the sum of the individual isotopes) can be better known (i.e. it can have a smaller variance) than each of the cross-sections of any of the individual isotopes. It follows that an experimental uncertainty in the order of 10% for a measurement of a natural sample does not imply that the cross-section for each of the individual isotopes is also known to approximately 10%.

In this section, we provide a demonstration of the consequences of reporting the cross-isotope correlations that are created by the evaluation of experimental data from natural isotopes. We consider the energy region of 30 to 45 eV for an experimental measurement of the capture cross-section using a natural gadolinium sample plotted in Figure 4.1. At first, we will only consider the effect of the statistical uncertainty on the experimental data. All of the systematic sources of uncertainty, such as uncertainties in the sample dimensions and experimental resolution function, are neglected in this demonstration.

In the energy region under consideration, 30 to 45 eV, only the isotopes ^{155}Gd, ^{156}Gd and ^{157}Gd have observed resonances. The statistical uncertainty from the experimental data is propagated to uncertainty in determining the resonance parameters for those three isotopes assuming no prior knowledge. The upper triangular portion of the joint resonance parameter correlation matrix is presented in Figure 4.2.

As argued above, due to resonance overlap between different isotopes, it is natural for correlations to arise between the resonance parameters of the three isotopes affecting the cross-section of the natural sample in this energy region. An error in the resonance parameters of one isotope will affect the certainty of the resonance parameters of the other isotopes. This is particularly evident in the appearance of the strong correlations in the upper left-hand corner of Figure 4.2, where the resonance parameters of ^{155}Gd are correlated to the resonance parameters of ^{157}Gd.

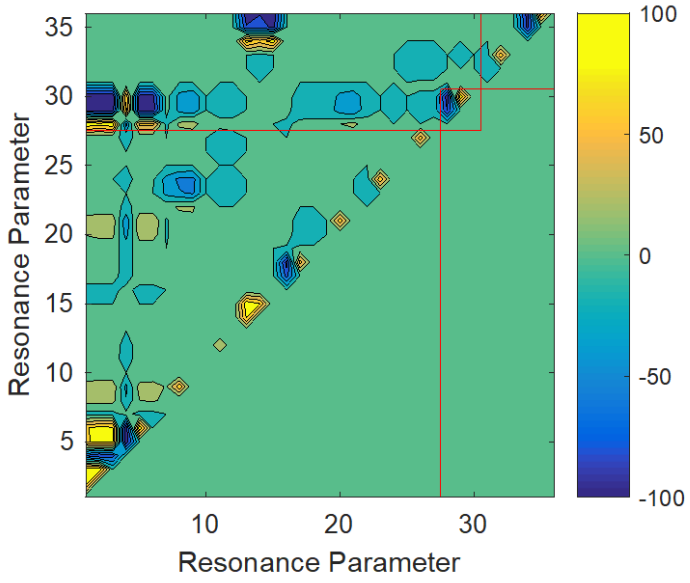
Figure 4.1. Experimental capture cross-section measurement based on a natural gadolinium sample



Source: ORNL, 2021.

Note that data are plotted with one standard deviation error bars arising solely from statistical uncertainty.

Figure 4.2. Upper triangular portion of the joint resonance parameter correlation matrix for resonances of ^{155}Gd, ^{156}Gd and ^{157}Gd



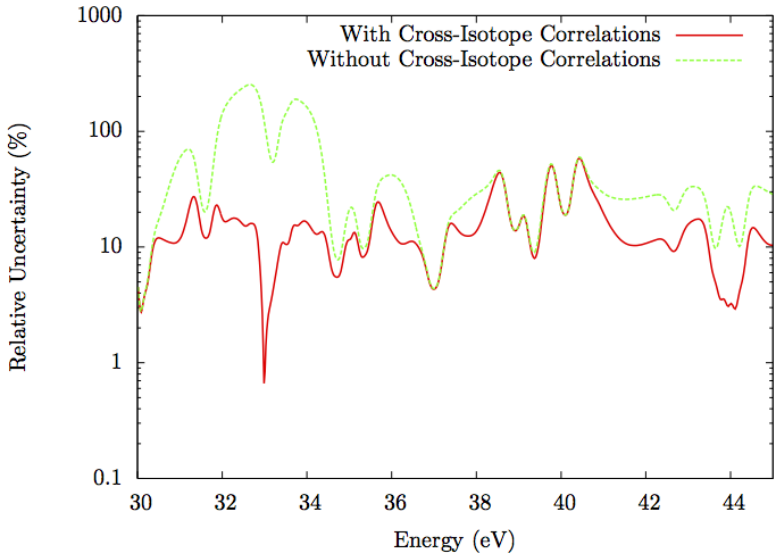
Source: ORNL, 2021.

Note that incident energy is between 30 and 45 eV and the scale is shown in percent.

Previously, however, cross-isotope correlations have not been reported in general-purpose nuclear data libraries. Figure 4.3 clearly demonstrates one of the consequences of neglecting the cross-isotope correlations. Figure 4.3 shows, in red, the relative uncertainty on the calculated cross-section that corresponds to the experimental measurement shown in Figure 4.1 as propagated from the full (with cross-isotope correlations) resonance parameter covariance matrix corresponding to Figure 4.2. The green curve in Figure 4.3 corresponds to the propagated uncertainty from the resonance parameter covariance matrix

if the cross-isotope correlations are neglected; only the correlations within the red squares of the individual isotopes are considered. The variance on each of the resonance parameters remains the same.

Figure 4.3. Relative uncertainty on the cross-section corresponding to Figure 4.1 propagated from the resonance parameter covariance matrix

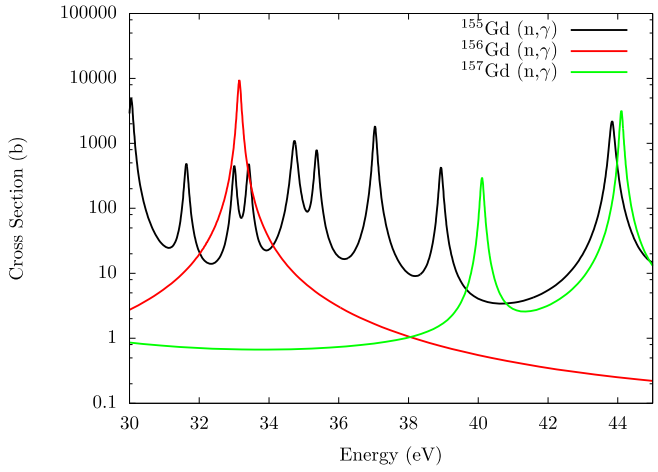


Source: ORNL, 2021.

Figure 4.3 shows that the systematic methodology of generalised linear least-squares results in a reasonable amount of uncertainty when the cross-isotope correlations are included. However, if the cross-isotope correlations are neglected, the propagated relative uncertainty jumps to rather large values at certain incident-neutron energies. In this case, it is evident that the cross-isotope correlations are such that the amount of uncertainty on the cross-section of a natural sample is reduced compared to the propagated uncertainty if it is assumed that the individual isotopes are uncorrelated.

Figure 4.4 shows the capture cross-section for the individual isotopes of gadolinium. ^{156}Gd only has one observed resonance in the energy region of 30 to 45 eV. However, that resonance dominates the cross-section of the natural sample around 33 eV, as shown in Figure 4.1. It is evident from Figure 4.3 that when the correlations between ^{156}Gd and the other isotopes are neglected, the propagated uncertainty in the cross-section of the natural sample increases significantly. The same effect is also observed when resonances ^{155}Gd and ^{157}Gd overlap around 44 eV.

Figure 4.4. Capture cross-section of the individual isotopes of gadolinium



Source: ORNL, 2021.

This is a contrived example which demonstrates the importance of including cross-isotope correlations when evaluating experimental measurements conducted on natural samples. Evaluators are encouraged to evaluate and report cross-isotope correlations when only natural samples are used as the basis for an evaluation. When highly enriched sample data are available, they will usually dominate the determination of the cross-section uncertainty for each isotope and cross-isotope correlations can be neglected.

4.1. Correlations in the thermal and intermediate range

In addition to the cross-correlations observed in the fast range, e.g. with the Jezebel PMF-1 benchmark, (Bauge and Rochman, 2018; de Saint Jean et al., 2015; de Saint Jean et al., 2018; Rochman et al., 2017), it is also possible to use experimental data in other energy ranges in order to create cross-correlations in the thermal and resonance ranges: within a specific isotope (e.g. nu-bar and (n,f) for ^{239}Pu), within elements (e.g. ^{235}U-^{238}U, or Si), and across elements (e.g. Si-O).

The method applied in the above references to calculate cross-section correlations is directly connected to the Bayesian Monte Carlo method (or BMC), but there are other very effective methods, such as GLLS. The first step is to produce so-called "random ENDF files", the second step is to use them to calculate a quantity (k_{eff}, group cross-section, dpa), leading to "random calculated quantities" and the last step is to extract weights for each random file, based on the level of agreement between the calculated and measured quantities.

The most commonly used experimental quantity is the k_{eff} of a specific critical system[2]. This is simply due to the availability of the benchmark models, of the possibility to produce processed nuclear data for the simulation code, and the limited required calculation time. One can also use other quantities from critical benchmarks, such as spectral indexes (Rochman et al., 2018a). In principle, any calculated quantity that can be compared to a

2. See, for example, "Correlation nu-sigma-chi in the fast neutron range via integral information" (Rochman et al., 2017) for PMF1, "Nuclear data correlation between different isotopes via integral information" (Rochman et al., 2018b) for IMF7, and "Monte Carlo nuclear data adjustment via integral information" (Rochman et al., 2018a) for a number of benchmarks considered simultaneously.

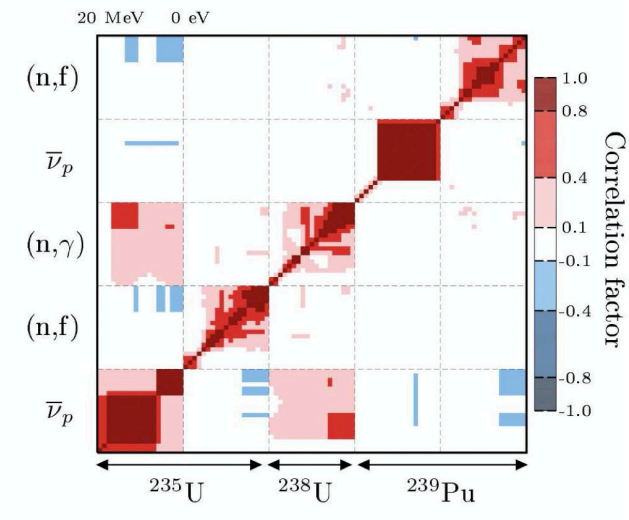
measured one can be updated (in a Bayesian sense), as well as its correlation matrix. If such experimental data are derived from a combination of calculated cross-sections, correlations between these cross-sections can be obtained. The work led by the IAEA (Sublet et al., 2019) presented an example where group cross-sections from an oxide component (SiO_2) are used to obtain cross-correlations between silicon isotopes and ^{16}O. As a final example, one can review "Correlation nu-sigma for U-Pu in the thermal and resonance neutron range via integral information" (Rochman et al., 2019), where the boron concentration from a specific reactor cycle is used to extract correlations between uranium and plutonium isotopes.

Two conditions still need to be met in the BMC approach: (1) the experimental data need to come with uncertainties which will encompass random calculated values (within one or two standard deviations), and (2) the experimental data need to be reasonably reproduced by the calculated average. Indeed, if the experimental uncertainties are very large compared to the spread of random calculated quantities, all the weights (for instance derived from χ^2 values) will be similar and no cross-correlations can be obtained. Additionally, if there is a large bias between the set of random calculated quantities and the measured one, all the weights will be very small, leading to no cross-correlations. Such issues have been detailed in some of the provided references.

From the cited studies, one important observation is that there is no general rule for cross-correlation. For instance, it is generally accepted that a k_{eff} value from a criticality benchmark will lead to an anti-correlation between the fission cross-section and the nu-bar of the most sensitive isotope (Bauge and Rochman, 2018; de Saint Jean et al., 2015; de Saint Jean et al., 2018; Rochman et al., 2017). One cannot nevertheless deduce that such cross-correlation is as general as a cross-section: it is obtained from the use of a specific type of experiment, and a different one may lead to different cross-correlations.

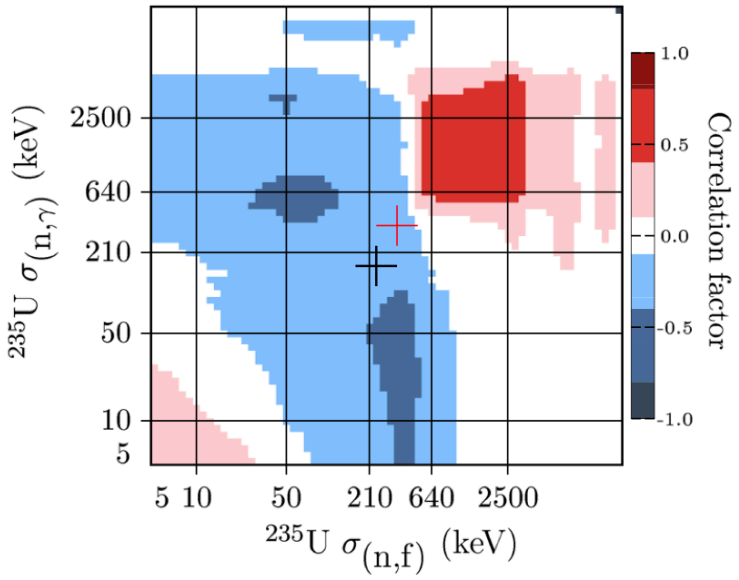
This is illustrated by the results presented in Figure 4.5, obtained from (Rochman et al., 2019). In (Rochman et al., 2018a). The IMF7 criticality benchmark and its k_{eff} are used to calculate various cross-correlations. The example presented here concerns the fission and capture cross-sections for ^{235}U. As observed, strong positive and negative correlations emerge. This is not the case for the same reactions when considering the thermal system of a pressurised water reactor (PWR) cycle with its measured boron concentrations.

Figure 4.5. Example correlations between isotopes due to the boron concentration in a PWR cycle



Source: PSI, 2021.

Figure 4.6. Correlations between two reactions of the same isotope using the Bigten benchmark



Source: PSI, 2021.

There are no strong correlations between these cross-sections, as observed in Figure 4.6. Therefore, correlations are only the reflection of the applied method and of the considered calculated and experimental data.

As a final comment, cross-correlations do not systematically come from integral experiments. A straightforward example is for experimental cross-sections: the vast majority of them are obtained relatively to a standard cross-section, implying a strong correlation between these quantities. Additionally, the use of oxide targets for cross-section measurements (as often used in EXFOR) will automatically lead to cross-correlations between oxygen and various target isotopes.

5. Intercomparison study

Four teams composed of the members of this subgroup participated in a computational intercomparison exercise on determining nuclear data correlations. This study looked at methods and techniques of retroactivity determining the sign and magnitude of correlation coefficients which are believed to be missing in evaluated nuclear data libraries. First and foremost, the methods developed in this study are not proposed as a long-term solution for addressing missing correlations from evaluated nuclear data libraries. The long-term solution, and a major recommendation of this subgroup, is for the nuclear data community to work together to provide the best estimates of all impactful correlations on evaluated nuclear data quantities during the generation of nuclear data evaluations and libraries.

The premise of the study undertaken here was that correlations are considered to be missing from currently published evaluated nuclear data libraries. The major source of those correlations was hypothesised to be the use of integral experiments, and in particular critical experiments, in the evaluation of the mean values in the library. Based on this knowledge, it may be possible to reconstruct the correlations which are not reported in today's nuclear data libraries. The objective of the study was to determine whether the sign and magnitude of those correlations can be consistently determined retroactively.

Four teams set to work using different initial nuclear data libraries and very different methodologies to estimate the missing correlations. The study was limited in scope to reconstructing the correlations between fission, capture and nu-bar for ^{239}Pu based on the scenario that a fast-spectrum plutonium-metal critical experiment was used as part of the evaluation of the central values for those quantities. The test was to see if the four groups, using different methods, would independently arrive at roughly the same values of the correlation coefficients. This is important because four independent efforts arriving at similar estimated values would give higher confidence in the validity of those values.

This report provides more than enough background information for the study described in this section. The details of the study, i.e. the assumptions and calculation methods used by each team and the quantitative results, are presented in (Sobes et al., 2021).

Overall, this computational intercomparison study found that the four participating teams, using different methods, could agree on the sign and approximate magnitude of the selected correlation coefficients. This validated the initial hypothesis that certain correlation coefficients that come from the use of criticality integral benchmarks in the evaluation can be reliably estimated after the fact. Furthermore, this study showed how those particular correlation coefficients can reduce the propagated uncertainty on similar systems and reduce the discrepancy between the observed spread in the C/E of mean values and the size of the error bars propagated from the current nuclear data covariance libraries.

The results of this intercomparison study, however, come with a caution not to over-extend the implications of the results. As has been stated above, this study was limited to correlations between fission, capture and nu-bar for ^{239}Pu based on a fast-spectrum plutonium-metal critical experiment. This set was particularly chosen because it was anticipated that 1) the correlations would be strong, 2) easily identifiable from the choice of integral experiment and 3) once identified, would have a significant impact on the propagated uncertainty to similar applications. Therefore, the caution is that this study does not suggest that all missing correlation coefficients from current nuclear data libraries can or should be estimated in this way. This study shows that in the absence of a better alternative, and only as a temporary remedy, some of the correlation coefficients can be

reliably estimated. The clear recommendation from this study and this subgroup is that nuclear data correlation coefficients need to be systematically evaluated and reported in future nuclear data library releases.

6. Formats and interpretation

This section contains general guidelines on how evaluated covariances stored in nuclear data libraries should be documented. Following these proposed guidelines should ensure that evaluated data and covariances are reproducible and are clearly interpretable by users. This text is based on the report in (CSEWG, 2012). It summarises guidelines concerning documentation, format, processing, verification, validation and associated tools which should be satisfied before nuclear data can be adopted by the Los Alamos National Laboratory (LANL) nuclear data team. These guidelines were established based on discussions between J.L. Conlin, M.G. Gray, A.P. McCartney, D. Neudecker, D.K. Parsons and M.C. White.

Here, we distinguish between requirements that "must", "ought" and "would be desirable" to be met by covariances in nuclear data libraries. Of course, we realise that many data and covariances currently in nuclear data libraries do not meet many of these documentation requirements. However, stating these requirements for new covariances might help improve the quality of documentation henceforth. Covariances, especially new ones, which do not satisfy the "must" requirements should be questioned by the community. Requirements that "ought" to be met concern information that is expected to be provided as best practice. However, we realise that often enough this information and associated tools might not be provided and the covariances might be adopted nonetheless. The requirements identified as "would be desirable" are recommended to be addressed by evaluators but covariances are unlikely to be rejected just because these requirements are not satisfied.

6.1. Documentation requirements

The following information should be available to document the sources of nuclear data and associated covariances:

- Evaluator(s): At least one corresponding author must be specified. This person should be able to answer questions future users might have regarding the covariances.
- Version: A number to identify the covariance evaluation unambiguously must be specified. This number can either be a version number or a commit number if versioning systems are used to track changes. If this is an update to a covariance matrix already existing in a library, the changes compared to the preceding library must be highlighted. An explanation for why this update is an improvement over the preceding version must be given and this explanation must be documented, either in the versioning system or another form of documentation (e.g. laboratory memo, report).
- Evaluation method: The algorithm used for the evaluation scheme ought to be specified, for example, Kalman filter, GLS, UMC-B (Brown et al., 2018). If specific input decks were used for the evaluation code, these along with the version number of the code ought to be provided.
- Nuclear model data and covariances: If a specific nuclear model was used as evaluation input, this model ought to be named. If a specific model code was used, it would be desirable to name it and provide its version. In order to guarantee reproducibility, it would be desirable that the initial model parameters and

uncertainties be provided along with any other information necessary to reconstruct the model data and covariances used for the evaluation. This could be done in the form of a journal publication that should be cited in MF=1.

- Experimental data and covariances: If experimental data were used in the evaluation process, these data ought to be documented by the number of their EXFOR-entry (Neudecker et al., 2018) or by citation to the associated references. Any changes of the original experimental data and their uncertainties undertaken as part of the evaluation process ought to be documented. In this case, the data as used for the evaluation ought to be provided either as a numbered memo or a laboratory report.

This information ought to be documented in journal publications, laboratory reports or memos with an unambiguously specified report or memo number. If the data are received by private communication, this communication must be at least documented in the versioning system. The journal article, report and/or memo number should be cited in the documentation section of the data set, e.g. in MF=1 of ENDF-6 formatted files (NEA, 2014).

It should be highlighted here that evaluated data and covariances for the same observable in a specific nuclear data library ought to be provided by the same evaluation procedure. The evaluated covariances can only yield realistic uncertainty information on mean values if the same experimental data, model, etc. were used to derive them.

Covariance format: The specific format the covariances are supposed to be given in (e.g. ENDF-6) must be clearly identified by citing unambiguously its manual. If the covariances are provided in a non-standard format or if no standardised one exists, it must be documented in detail in an easily retrievable form. The same is true if small changes compared to the standard format were undertaken. Otherwise, the covariances could be incorrectly interpreted by users.

Covariance processing and verification: The processing and verification of covariances ought to be documented in versioning systems, laboratory reports, etc., along with the version number of the codes used for this task. It would be desirable to document the input files for running these codes, whenever possible.

Of course, it would be desirable that all covariances be processed before they are accepted in new nuclear data libraries. However, whenever new covariance formats come along (e.g. currently for fission yields and thermal scattering law covariances), it takes some time until new processing capabilities are developed.

For the *validation and quality assessment of the covariances*, the following information ought to be documented:

- Validator(s): One person serving as point of contact for questions regarding the quality of the covariances would be desirable to be named. We understand that validation and quality assessment are often not comprehensively undertaken and that a point of contact might not be available for all covariances.
- Quality assessment and validation of covariances: If a validation or quality assessment of the covariances was undertaken, the results ought to be documented either as a report or in the versioning system. It would be especially important for users of these covariances to know whether the uncertainties are reliable, under- or over-estimated.

The main information for documentation can be separated into (1) critical, (2) essential and (3) desirable. These are summarised in Table 6.1.

Table 6.1. Summary of the documentation guidelines

Data type	Requirement level	Example
Evaluator(s)	Critical	Name, institution, e-mail and/or contact details
Version	Critical	Version number of the evaluation, commit number and/or hash for relevant version control systems, indicative message
Evaluation method	Essential	Description of the methodology employed (e.g. generalised least-squares using GMA version X)
Nuclear model data and covariances	Essential	Description of the models employed, versions, parameters or other relevant data (e.g. CoH version X as described in publication Y with parameters Z)
Experimental data and covariances	Essential	Experimental data with identifiers, typically given with EXFOR compilation numbers and/or publication references, including any adjustments
Covariance format	Critical	Data format used (e.g. ENDF-6) including any relevant version number or modifications
Covariance processing and verification	Essential	Codes, versions and descriptions of processes used in the processing of the data (e.g. NJOY version X using Y models)
Validator(s)	Desirable	Name, institution, e-mail and/or contact details
Quality assessment and validation of covariances	Essential	Description of the quality assessment process and references to any applicable documentation

Source: LANL, 2021.

6.2. Verification of eigenvalues and decompositions

One common application for nuclear data covariance matrices is generating random samples of data for use in uncertainty quantification (UQ) studies. There are several methods for generating these random samples. For example, a covariance matrix C can be Cholesky-decomposed to obtain a triangular matrix M such that

$$M \times M^T = C$$

By drawing a random (normally-distributed) vector of numbers V, a new realisation can then be generated by simply multiplying the triangular M matrix and V. Alternately, covariance matrix C can be diagonalised to obtain eigenvalues λ and eigenvectors Λ such that

$$C = \Lambda \lambda \Lambda^T$$

Then the realisation R can be constructed by:

$$R_i = \sum_j \sqrt{\lambda_j V_j \Lambda_{j.i}}$$

where Vj is again a normally-distributed random value. Unfortunately, nuclear data covariance matrices are frequently not positive-definite and both of these techniques have trouble dealing with non-positive-definite matrices. Nuclear data covariance matrices may be positive semi-definite (with many eigenvalues near 0), but more often they include small negative eigenvalues, especially when (due to limitations in the ENDF-6 format) covariance matrices are stored using only 8-9 significant digits.

Since nuclear data covariances are often dominated by only a few non-negligible eigenvalues, it can be advantageous to disregard small eigenvalues (both positive and negative) and only store the principal components. For example, the ENDF-VIII neutron sub-library prompt fission neutron spectrum covariance for incident neutrons between 5 and 6.5 MeV is a 591x591 matrix, but the matrix is dominated by the six largest principal components (these are sufficient to reproduce all matrix elements to within an absolute tolerance of 10^{-8} and a relative tolerance of 10^{-5}). Storing only the principal components saves space and avoids an extra step when sampling since the matrix is already diagonalised.

A recent proposal would extend the GNDS format to support storing matrices in "sandwich product" form, i.e. as the matrix product $A \times b \times A^T$. This format is meant to be flexible enough to handle both diagonalised matrices (where b is diagonal) and more general parameter covariances with sensitivity matrices for rows and columns.

6.3. Thermal scattering law (TSL) covariances in GNDS

Recent improvements in both experiments and modelling have led to a renaissance in thermal scattering law (TSL) evaluations, relevant for describing coherent and incoherent scattering reactions between low-energy neutrons and target materials where molecular effects must be taken into account. No current TSL evaluations include covariances, but an effort is underway to produce and store TSL covariances. This effort has initially focused on adding covariances to incoherent inelastic scattering reactions, where the double-differential cross-section is parameterised as a function of momentum transfer α, energy transfer β and temperature T, denoted:

$$S = S(\alpha, \beta, T)$$

Since S is a function of three independent variables, the full covariance matrix is in principle a 6-dimensional array correlating each combination of α_i, α_j, β_i, β_j, T_i and Tj.

GNDS supports storing higher-dimensional arrays, but the current expectation is that rather than storing a 6-dimensional covariance, evaluators will likely instead break it into multiple 2-dimensional matrices. Each of these matrices would store the covariance of S with respect to $\alpha_{i,j}$, and each would correspond to a given range of β and a given temperature T[3]. Breaking the full covariance matrix up into these sub-sections provides more flexibility compared to storing a 6-dimensional array: different α grids may be used for different combinations of β and T rather than enforcing a uniform grid on the entire covariance.

Covariances for coherent and incoherent elastic scattering are also possible. These are simpler to support since the parameterised forms have fewer independent variables: for coherent elastic the cumulative structure factor is tabulated as a function of incident energy and temperature, and for incoherent elastic the Debye-Waller integral is a function of temperature only.

3. Note that this decomposition of the full covariance into multiple sections is similar to how ENDF treats covariances for outgoing energy and/or angle distributions.

6.4. Prompt fission neutron spectra correlations

It should be highlighted that measured prompt fission neutron spectra (PFNS), fission cross-section ((n,f) cross-section) and average prompt neutron multiplicities (<νp>) are all related to the same physics reaction (i.e. fission). This becomes obvious when one looks at the neutron transport equation to simulate k_{eff}. These three nuclear data observables form together the fission neutron source term.

However, the ENDF-6 format (CSEWG, 2012) does not make it possible to store any correlations between uncertainties of these three nuclear data observables. One might argue that this is strictly speaking not necessary as the PFNS, (n,f) cross-section and $<\nu_p>$ are often evaluated separately. To give an example: the ^{239}Pu(n,f) cross-section in ENDF/B-VIII.0 (Brown et al., 2018) was evaluated based on mostly experimental data as part of the Neutron Data Standard project (Neudecker et al., 2018), the PFNS was evaluated with the Los Alamos model and ($<\nu_p>$) based on experimental data in the fast range and partially adopted from (NEA, 2014). All these three evaluations were independently undertaken by separate groups of experts, model, codes and experimental data. Hence, the resulting evaluated data can be treated as independent and the covariances between the different observables are indeed zero.

However, with the advent of advanced fission modelling codes such as (Vogt et al., 2009; Litaize and Serot, 2009; Becker et al., 2013), the same codes are able to yield all observables simultaneously based on a joint set of model parameters. Usage of these new codes for nuclear data evaluations would certainly introduce correlations between PFNS, (n,f) cross-section and $<\nu_p>$ uncertainties that should be provided in nuclear data libraries. In addition to that, adjustment of data and covariances with respect to k_{eff} would lead also to correlations between uncertainties of PFNS, (n,f) cross-section and $<\nu_p>$.

Hence, formats should be developed to store correlations between PFNS, (n,f) cross-section and $<\nu_p>$ in nuclear data libraries. Along the same lines, formats should be developed to store correlations between PFNS uncertainties of different isotopes and incident energy groups. Currently, ENDF-6 format does not allow them to be stored. Hence, they are implicitly zero. Work by Rising et al. published in the International Criticality Safety Benchmark Evaluation Project (NEA, 2019a) showed that neglecting correlations between PFNS uncertainties of different isotopes can lead to a reduction of k_{eff} uncertainties of selected critical assemblies (Jezebel-240, Flattop-25, Big Ten) stemming from the PFNS by ~5-20%.

The same reference showed that the k_{eff} uncertainties stemming from the PFNS were underestimated by more than 50% for Jezebel, Flattop-Pu and Thor critical assemblies if the PFNS correlation matrices were stored for each incident-neutron energy the PFNS is given and implicitly assumed 0 between them. However, the ENDF-6 (CSEWG, 2012) PFNS covariance format makes it possible to use covariances defined for a specific incident-neutron energy range. The correlations between PFNS uncertainties at different incident-neutron energies are very similar for the same physics range, i.e. for first chance fission, second-chance fission, third-chance fission, etc., because the underlying physics of the model is the same. That can be exploited by giving a covariance for one physics range.

This reasoning was applied for providing covariances grouped by a physics-defined incident-neutron energy for ^{235}U and ^{239}Pu for ENDF/B-VIII.0 (Brown et al., 2018; Neudecker et al., 2018). The impact of this grouping was studied by propagating these grouped covariances versus covariances considering correlations between PFNS uncertainties at all incident-neutron energies to obtain k_{eff} uncertainties of Jezebel due to the PFNS. The former k_{eff} uncertainty was -0.36 % lower compared to the k_{eff} uncertainties

caused by the PFNS using the full covariance matrix. The Godiva k_{eff} uncertainties due to the PFNS increase by a negligible 0.47 pcm due to grouping the PFNS. That is an increase of 0.91 % in the k_{eff} uncertainty due to the grouping. So, in short, with the current ENDF-6 format, one can provide incident-neutron energy grouped covariances that retain the correct uncertainties in k_{eff} of most critical assemblies for single isotopes. Godiva and Jezebel were used as examples as those are benchmarks among those with the fastest spectra of ICSBEP critical assemblies (NEA, 2019a). Thus, they are most sensitive, but still very little so, to second-chance fission processes.

When one samples from these incident-neutron energy range grouped PFNS covariances, one can still reproduce the original PFNS well except at high outgoing-neutron energies. There, the tail changes with incident-neutron energy. However, the PFNS values are small above 10 MeV outgoing-neutron energy and critical assemblies are usually not sensitive to these energy ranges.

Hence, in short k_{eff} uncertainties should not be impacted by the missing format to allow storing correlations between PFNS uncertainties at different incident-neutron energies. However, this should be tested for other types of application calculations that are sensitive to broader incident-neutron energy ranges.

6.5. Data format for angular distributions

In the ENDF-6 format (CSEWG, 2012), angular distributions are expressed as normalised probability distributions $\int_{-1}^{1} f(\mu, E)d\mu = 1$ where $f(\mu, E)d\mu$ is the probability that a particle with incident energy E is scattered in the interval $d\mu$ around an angle with cosine μ. Then, under the general assumption of azimuthal symmetry over the scattering angle, the format allows representing the distribution in terms of NL (≤ 64) Legendre polynomials,

$$f(\mu, E) = \sum_{l=0}^{NL} \frac{2l+1}{2} a_l(E) P_l(\mu)$$

where the energy and angular dependence is decoupled. Here, P_l is the l-th order Legendre polynomial and a_l its corresponding coefficient. Generally, NL increases with incident energy as the angular distribution becomes more forward-oriented. Alternatively, $f(\mu, E)$ can be tabulated as a function of the incident energy and angle cosine.

6.5.1. Angular distribution covariance matrices

Covariance matrices for the angular distribution of secondary particles emitted from neutron-induced reactions, such as elastic and inelastic scattered neutrons, became available in the early 1990s in the EFF-2 evaluations. The matrices were prepared in terms of covariances between Legendre polynomial coefficients and were reported into section MF=34 of ENDF-6 formatted nuclear data files. Covariance terms up to P6 were evaluated for the elastic scattering cross-section of ^{56}Fe, ^{52}Cr, ^{58}Ni and ^{60}Ni (Kodeli, 2018). More recently, the JENDL-4.0 (Shibata et al., 2011) and ENDF/B-VII.1 (Chadwick et al., 2011) nuclear data libraries included covariance matrices for P1 for isotopes of Fe, U, Pu and others. The CIELO evaluations for ^{235}U and ^{238}U of the WPEC Subgroup 40 project (NEA, 2019b) were adopted in the ENDF/B VIII.0 library (Brown et al., 2018). They include covariances of the Legendre coefficients of the elastic angular distributions in the laboratory co-ordinate system up to order P2. They also include cross-correlation with the elastic cross-section, as well as between the Legendre coefficients of different order. The covariance matrix prior was generated by Monte Carlo sampling of the nuclear model parameters. In the assembled global covariance matrix the scattering Legendre moments

were included. In a subsequent generalised least-squares fitting of experimental data for other reaction cross-sections, the covariances of the scattering moments were automatically adjusted consistently through correlations. The covariances of the Legendre coefficients were reconstructed from the scattering moments at the ENDF-6 formatting stage. This approach has the advantage that scattering moments which appear in deterministic transport equations are as accurate as possible. Unfortunately, processing codes like NJOY were not upgraded to take this form of covariance information with correlations into account. Even more covariances for angular distributions were produced with the TASMAN code and included in the TENDL libraries for both elastic and inelastic scattering cross-sections (Koning et al., 2019). These covariances extend up to terms higher than the first Legendre polynomial. The JEFF-3.3 library recently adopted these covariances for several of their nuclide evaluations (Plompen et al., 2020). The importance of such covariances was reported for neutron shielding application (Jouanne, 2014; Kos et al., 2018).

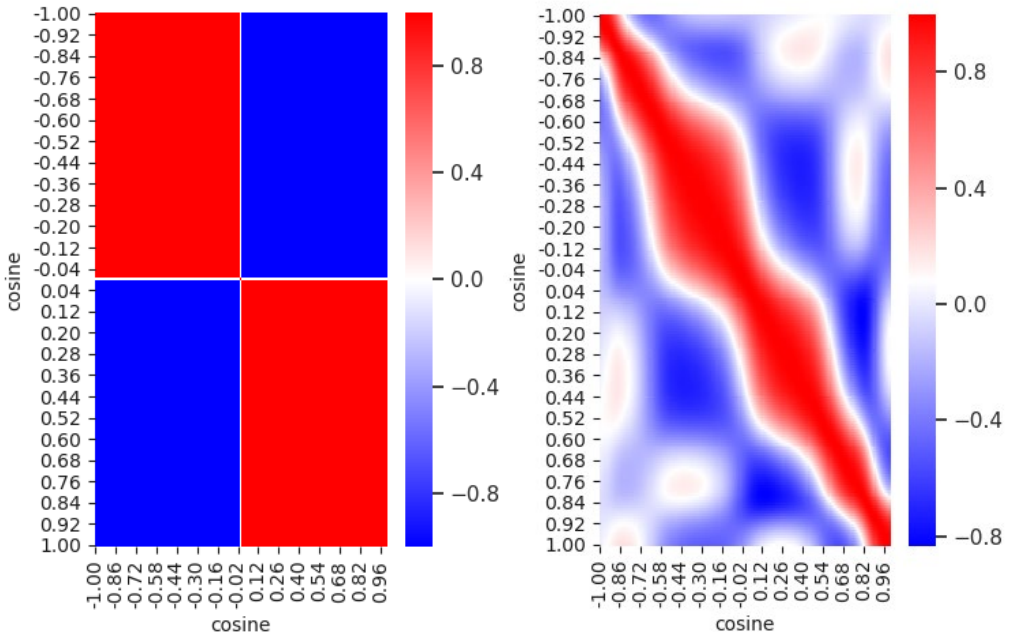
The first order Legendre coefficient in the laboratory system is by definition the average cosine of scattering (known as "mu-bar"), which can be represented in ENDF by MF=3/MT-251, making the processing of covariances equal to other cross-sections in MF=33. However, this approach is limited to the order P1. Computer codes such as NJOY (MacFarlane et al., 2019) have since developed capabilities to process energy-dependent covariances for P1 in MF=34 into multi-group formats, which can be subsequently used for linear uncertainty propagation by codes such as SUSD3D (Kodeli and Slavič, 2017), but the coding is not in place to process cross-correlations. On the other hand, nuclear data sampling codes such as SANDY (Fiorito et al., 2017) can technically be used to test and propagate covariance data for any polynomial order.

SANDY is a nuclear data sampling tool that, for a given evaluated nuclear data and covariance file formatted according to the standard ENDF-6 rules, can produce perturbed copies of the file that statistically reflect the covariance information. The perturbed files are generated by applying random perturbation coefficients sampled from the multivariate Gaussian distribution defined by the evaluated data. SANDY can perturb any type of nuclear data for which covariances are provided, including cross-sections and fission neutron multiplicities, as well as emitted energy and/or angular distributions. SANDY enforces consistency with the conservation rules, e.g. by renormalising the probability distribution functions. The perturbed files are suitable for Monte Carlo uncertainty propagation calculations, where a given model is solved repetitively, each time adopting a different file. From the statistical analysis of the resulting output predictions one can calculate the distributions of any calculated quantity, including their mean, variance and any other moment.

In the recent work of Fiorito et al. (2019), the ICSBEP benchmark PU-MET-FAST-006 (NEA, 2019a) was selected for the large sensitivity of its k_{eff} to the angular distribution of the neutrons scattered in ^{238}U and out of the system. The uncertainty contribution of the angular distribution was quantified using SANDY and the covariance information for the Legendre polynomial coefficients up to the 6th order taken from the ^{238}U evaluation in JEFF-3.3. The resulting k_{eff} distribution had a standard deviation of about 2 000 pcm and was significantly non-Gaussian, highlighting a non-linear model. The uncertainty breakdown showed that for such benchmarks with heavy reflectors the P1 term describes only a fraction of the collective uncertainty coming from all terms of the elastic scattering angular distribution. Also, it was highlighted that using a Legendre polynomial representation to store the covariance information for angular distributions as a function of the scattering angle implicitly imposes correlations between scattering angles that are merely an artefact of the mathematical definition of the polynomials. In particular, Figure 6.1 shows that the P1 term alone generates a strong anti-correlation between the forward and backward angular directions as a reflection of the shape of the first polynomial

term. This effect can be avoided by adopting higher order polynomials for reporting the covariance data of angular distributions, or alternatively, by adapting the format to accept tabulated covariances as a function of the incident energy and scattering cosine.

Figure 6.1. JEFF-3.3 ^{238}U elastic scattering angular correlation matrices for an incident-neutron with energy 1.8 MeV



a) covariances for P1 only

b) covariances up to P6

Source: SCK·CEN, 2021.

Note: The covariance matrices for Legendre polynomial coefficients are shown up to first order (P1) on the left and up to sixth order (P6) on the right. These were converted into covariances between scattering angles with cosines listed on the axes. The colour axis represents the probability distribution.

6.6. Covariance information of double-differential data

The formats for the covariance information of secondary particle correlated energy/angle distributions are not defined. While energy/angle correlations need to be considered explicitly in transport calculations, especially at incident energies above several MeV, the uncertainty estimates are considered less sensitive to the exact representation. For the purpose of calculating relative uncertainties it is probably sufficient to split double-differential data separately into angular distributions in MF=4 and energy distributions for the same reaction in MF=5. The covariances can then be represented in the conventional way in MF=34 and MF=35 in ENDF terminology. This procedure has not been exploited so far, but it is worth exploring since uncertainties in the particle emission spectra and the corresponding angular distributions certainly contribute to the uncertainties in the integral results.

6.7. Total Monte Carlo (TMC) versus covariance data

Covariance data are added to nuclear data libraries to propagate correlated uncertainties from fundamental physics to nuclear applications. Many of the covariance data in nuclear data libraries, especially in MF33 and MF34, are induced by more fundamental uncertainties from e.g. nuclear model parameters which end up in large covariance matrices for cross-sections, angular distributions etc. accompanying the final results.

There is a different method for uncertainty quantification and propagation and that is to perform a statistical sampling of only the basic nuclear data uncertainties and to produce a statistical ensemble of nuclear data libraries out of that. This method was advocated in 2008 (Koning and Rochman, 2008), and later coined by the Head of the US National Nuclear Data Center, Michal Herman, as "Total Monte Carlo" (TMC).

In the case of a general-purpose nuclear data library, as fundamental uncertainty data one would consider:

- resonance parameters and their correlated uncertainties;
- EXFOR data and their correlated uncertainties;
- nuclear model parameters and their correlated uncertainties.

One could draw one sample from all these probability distributions, using correlated sampling, perform a nuclear model calculation for the part which requires theory (the fast energy range) and produce one complete "random" nuclear data library out of that. The term "random" is one of the most unfortunate adjectives adopted in science: in common language it means a "shot in the dark", i.e. a completely wild guess as if the users do not know what they are doing. Instead, in the present context it is meant as a carefully sampled deviation from the average inside a known uncertainty band.

Next, one can process this collection or random files that form a so-called 'random library' and do an applied calculation, e.g. a k_{eff} estimate with the Monte Carlo N-Particle (MCNP) code, for that particular sample. This process can then be repeated for say 100 to 1 000 random samples, and uncertainties and covariances for the final results then become available, in the form of e.g. a probability distribution for k_{eff}. This could then be done for the entire ICSBEP suite of benchmarks (Rochman et al., 2009).

Moreover, covariance matrices for the contents of the nuclear data library are no longer needed, although they can easily be stored in separates file in the process for consultation, since all correlations are intrinsic in the entire ensemble of nuclear data libraries, where they are reflected for all energies and reaction channels and data types, including cross-correlations.

As discussed in many publications since 2008, there are disadvantages and advantages to TMC. The two main disadvantages are:

- The calculation time: e.g. an MCNP calculation has to be performed many times instead of only once. An effort to mitigate this has been done with so-called "Fast TMC" (Rochman et al., 2014).
- The evaluation process needs to be extremely reproducible, and when one insists on starting from the three fundamental data classes listed above, only one system is currently able to do this: the so-called T6 system built around TALYS, which also produces the TENDL library. In the past decade also "semi-TMC" has been done (Diez et al., 2015; Fiorito et al., 2017), which entails producing random

nuclear data libraries by sampling from existing nuclear data libraries which contain covariance data.

There are various advantages for TMC as well. A few important ones are:

- Both the uncertainty quantification and propagation can be considered as more exact since all correlated uncertainties are taken into account, and not only those that are limited by the ENDF format. This has been shown through comparison of the TMC method with the method using covariance data (Rochman et al., 2011).
- A covariance file cannot describe higher moments without changes to current formats and as a result any features that could be represented with higher-order moments (e.g. skewness) cannot be described. Covariance matrices also require an energy discretisation that may lead to non-trivial loss of fidelity.
- Covariance matrices do not need to be dragged along in large computational schemes which require covariance matrices, possibly in various grouped representations and mixed with further processing steps. For example, TMC has even been applied to a control rod ejection case where a statistical ensemble of nuclear data libraries was propagated all the way to thermohydraulic calculations with the PANTHER code (da Cruz et al., 2014a) and fuel inventory of a PWR assembly (da Cruz et al., 2014b).
- TMC opens up uncertainty propagation for classes of nuclear data for which either no format exists or for which the covariance representation would be too complicated.
- Propagating the covariance data to macroscopic system parameters (e.g. k_{eff}) requires the system sensitivity matrix, which is not always trivial to obtain. The sensitivity matrix can also be computationally costly to produce and assumes linearity of the system, an assumption that is not always well-motivated. Higher moments of the system can be obtained, but this is even more challenging and, independent of the truncation, some moments of the system response will always be lost.

Especially the last item in the list above connects well to this report. If we consider a complete general-purpose nuclear data library, i.e. one which extends from MF1 to MF15 in ENDF jargon, it is interesting to see which parts of the library can be accompanied with covariance data. If not, this is then an advantage for TMC, since a statistical ensemble of e.g. 1 000 libraries which all range from MF1 to MF15 automatically contain the entire correlated probability distributions for all data inside. TENDL-2019 can be considered a complete covariance nuclear data library, in the sense that it contains covariance as much as the ENDF format reasonably allows. Therefore, the list below relates to TENDL. If we go over the data library class by class, we find the following:

- MF1: Covariance data for nu-bar are essential for integral use of data libraries, and an appropriate MF31 representation exists.
- MF2: Covariance data for resonance parameters have always been considered as essential and a MF32 taking the required correlations into account exists. A so-called “compact MF32” format can be used for very large covariance matrices. Several nuclear data users prefer to have the covariance data for the resonance range pointwise in MF33.
- MF3: The ENDF format allows for both inter- and intra-channel covariance data for all cross-sections.

- MF4: Covariance data are used for elastic scattering angular distributions in MF34/MT2.
- MF5: Covariance data are used for fission neutron spectra, prompt in MF5/MT18, and delayed in MF5/MT455.
- MF6: MF36 does not, and very probably will not, exist. This means one can generally not find covariance data for single- and double-differential emission spectra, exclusive gamma-ray branching ratios, residual production cross-sections and recoils. There is a theoretical possibility to represent the covariance data of MF6 single-differential emission spectra in MF35, and to store covariance data for the MF6 residual production cross-sections in MF40, but this has seldom been tried and the capabilities for these data to be processed are not guaranteed.
- MF7: Thermal scattering data. MF37 does not (yet) exist and there is no accepted formalism within the ENDF-6 format, though efforts in SG-48 are underway to tackle this. The most mature proposal requires the new GNDS format (NEA, 2020b). Meanwhile, thermal scattering using TMC has been reported in (Rochman and Koning, 2012).
- MF8: Fission yield data. Only variances are taken into account in a specially adapted format for MF8, while a full covariance format is under discussion. For a TMC fission yield library this is no issue. For decay data, the same principle applies, only for the most basic quantities are variances provided, while in a TMC approach one would generate randomly sampled decay data libraries with all generated parameters sampled from the distributions defined by the uncertainty bands.
- MF9/10: Covariance data for isomeric production can be represented in MF40 as in MF33 for cross-sections.
- MF12: Gamma branching ratios and yields, MF42 does and will not exist, uncertainty quantification/propagation is only possible with TMC.
- MF14/15: Gamma secondary distributions, MF44/45 does and will not exist, uncertainty quantification/propagation is only possible with TMC.

In summary, complete covariance data, i.e. including all required cross-correlations, is not available, or very difficult to represent, for thermal scattering, fission yield data, decay data, single- and double-differential spectra, residual production cross-sections, recoils, gamma-ray production cross-sections or branching ratios.

Alternatively, if one were to generate a statistical ensemble of say random nuclear data libraries, requiring perhaps 1 000 samples or more, this could be used to quantify and propagate uncertainties for:

- general-purpose data library;
- fission yield data library;
- thermal scattering data library;
- decay data library.

All correlated uncertainty data that one would have available at the start of the evaluation process from e.g. experimental data, standards evaluations or nuclear model parameters would exactly be propagated to applications.

7. Conclusions and recommendations

This report has compiled and presented a summary of the work of this subgroup. This report certainly is not the final word on nuclear data covariance. The members of SG-44 rather hope that through their work and this report they will accomplish three major objectives. First, they hope to provide a comprehensive review of nuclear data covariance activities and research interests around the world in recent years. Second, they hope to educate the reader about the importance of nuclear data covariance and about the current recommended practices. Third, they hope to have correctly identified and communicated the most important needs of the nuclear data community with respect to future work on nuclear data covariance. Rather than providing the final word on nuclear data covariance, this subgroup hopes to revitalise the conversation on this important topic.

A fundamental challenge in nuclear data covariance evaluation still remains. Nuclear data covariance is a statement of confidence in the evaluated central value and cannot be experimentally tested or validated. However, through the work of members of SG-44, this report presented some recommended "sanity-checks" on reported nuclear data covariance, which is great progress in the field.

This subgroup on covariance data evaluation takes no unified stand on the question of "Whether integral experiment information should be systematically included in the nuclear data evaluation step or only reserved for validation efforts." In fact, the members of the subgroup had different perspectives on this question, which led to vigorous discussions in almost every meeting of the subgroup, but no consensus. The subgroup believes it is important to state this clearly to the reader and to future generations of scientists.

On the other hand, all of the members of the subgroup agreed that much more rigorous documentation of the covariance evaluation process is needed. This is the most significant recommendation of this subgroup. New nuclear data evaluations need to describe in rigorous detail the evaluation process, whether and how integral experiment information was used and how covariance was evaluated.

Many nuclear data covariance needs have been discussed. However, the exact prioritisation of the needs is rather application specific. A number of major areas of nuclear data were identified where covariance evaluation was missing or was only in its beginnings. These areas include:

- cross-correlations between different reactions and isotopes which arise if integral experiments are used in the evaluation of nuclear data;
- initial angular distribution covariance data have been generated for only a few isotopes; more angular distribution covariance data is needed as well as robust methods to propagate the uncertainty through to applications;
- thermal scattering law covariance data is entirely missing from current evaluated nuclear data libraries; this subgroup has worked with Subgroup 42 and hopes to have inspired current Subgroup 48 to seriously consider the generation of thermal scattering law covariance data and methodologies for its propagation to applications.

Lastly, this subgroup would like to highlight the excellent work in generating the General Nuclear Database Structure. This format will provide the much-needed flexibility to store covariance data. This subgroup highly recommends that users and producers of nuclear data covariance adopt the new format as quickly as possible.

While the term of this subgroup is complete, the work on nuclear data covariance lives on in several activities under the Working Party for Nuclear Data Evaluation Co-operation. With the strongest recommendation of this subgroup for careful documentation of nuclear data covariance evaluation, one potential solution is presented by the efforts of WPEC Subgroup 49 on Reproducibility in Nuclear Data Evaluation. A complete, end-to-end computer code evaluation system would provide the necessary reproducibility, traceability and documentation that is missing so much in current nuclear data evaluation processes.

References

Archier, P., C. De Saint Jean, O. Litaize, G. Noguère, L. Berge, E. Privas and P. Tamagno (2014a), "CONRAD evaluation code: Development status and perspectives," *Nuclear Data Sheets* 118, pp. 488-490, DOI:10.1016/j.nds.2014.04.114.

Archier, P., G. Noguère, C. De Saint Jean, A. Plompen and C. Rouki (2014b), "New JEFF-3.2 sodium neutron induced cross-sections evaluation for neutron fast reactors applications: from 0 to 20 MeV," *Nuclear Data Sheets* 118, pp. 140-143, DOI: 10.1016/j.nds.2014.04.020.

Bauge, E. and D. Rochman (2018), "Cross-observables and cross-isotopes correlations in nuclear data from integral constraints", *European Physical Journal* N, Vol. 4, pp. 35, DOI:10.1051/epjn/2018011.

Becker, B., P. Talou, T. Kawano, Y. Danon and I. Stetcu (2013), "Monte Carlo Hauser-Feshbach predictions of prompt fission γ rays: Application to $n_{th}+^{235}U$, $n_{th}+^{239}Pu$, and ^{252}Cf (sf)", *Physical Review* C 80, 044611, DOI:10.1103/PhysRevC.87.014617.

Boldeman, J.W. and J. Fréhaut (1980), "The foil thickness correction in ν_{tot} measurements and the (ν-η) discrepancy", *Nuclear Science and Engineering*, Vol. 76, 1, pp. 49-52, DOI:10.13182/NSE80-A19292.

Brown, D.A. et al. (2018), "ENDF/B-VIII.0: The 8th major release of the nuclear reaction data library with CIELO-project cross sections, new standards and thermal scattering data", *Nuclear Data Sheets*, Vol. 148, pp. 1-142, DOI:10.1016/j.nds.2018.02.001.

Capote, R. and D. Smith (2008), "An investigation of the performance of the unified Monte Carlo method of neutron cross section data evaluation," *Nuclear Data Sheets* 109, 2768 DOI:10.1016/j.nds.2008.11.007.

Capote, R. et al. (2020), "Unrecognized sources of uncertainties (USU) in experimental nuclear data", *Nuclear Data Sheets*, Vol. 163, pp. 191-227, DOI:10.1016/j.nds.2019.12.004.

Carlson, A.D. et al. (2018), "Evaluation of the neutron data standards", *Nuclear Data Sheets*, Vol. 148, pp. 143-188, DOI:10.1016/j.nds.2018.02.002.

Chadwick, M.B. et al. (2018), "CIELO Collaboration Summary Results: International evaluations of neutron reactions on uranium, plutonium, iron, oxygen and hydrogen", *Nuclear Data Sheets*, Vol. 148, pp. 189-213, DOI:10.1016/j.nds.2018.02.003.

Chadwick, M.B. et al. (2011), "ENDF/B-VII.1 Nuclear data for science and technology: Cross sections, covariances, fission product yields and decay data", *Nuclear Data Sheets*, Vol. 112, pp. 2887-2996, DOI:10.1016/j.nds.2011.11.002.

CSEWG (2012), ENDF-6 Formats Manual, Data Formats and Procedures for the Evaluated Nuclear Data Files ENDF/B-VI and ENDF/B-VII, edited by A. Trkov, M. Hearman and D. Brown, CSEWG Document ENDF-102, Report BNL-90365-2012 Rev. 2, National Nuclear Data Center, Brookhaven National Laboratory.

da Druz, D.F., D. Rochman and A. Koning (2014a), "Propagation of nuclear data uncertainty for a control ejection accident using the Total Monte Carlo method", *Proceedings of PHYSOR 2014* – The Role of Reactor Physics Toward a Sustainable Future, Kyoto, Japan, 28 September – 3 October 2014.

da Cruz, D.F., D. Rochman and A. Koning (2014b), "Uncertainty analysis on reactivity and discharged inventory due to $^{235,238}U$, $^{239,240,241}Pu$ and fission products nuclear data uncertainties - application to a pressurized water reactor fuel assembly", *Nuclear Technology*, Vol. 185, 2, pp. 174-191, DOI:10.13182/NT12-154.

Diez, C.J., O. Buss, A. Hoefer, D. Porsch and O. Cabellos (2015), "Comparison of nuclear data uncertainty propagation methodologies for PWR burn-up simulations", *Annals of Nuclear Energy*, Vol. 77, pp. 101-114, DOI:10.1016/j.anucene.2014.10.022.

de Saint Jean, C., G. Noguère, B. Habert and B. Iooss (2009), "A Monte Carlo approach to nuclear model parameter uncertainties propagation", *Nuclear Science & Engineering*, Vol. 161, pp. 363-370, DOI:10.13182/NSE161-363.

de Saint Jean, C., G. Habert, P. Archier, G. Noguère, D. Bernard, J. Tommasi and P. Blaise (2010), "Uncertainty evaluation of nuclear reaction model parameters using integral and microscopic measurements with the CONRAD code", *Journal of Korean Physics Society*, Vol. 59, pp. 1276–1279, DOI:10.1051/epjconf/20100804002.

de Saint Jean, C., P. Archier, E. Privas, G. Noguère, O. Litaize and P. Leconte (2015), "Evaluation of cross section uncertainties using physical constraints: Focus on integral experiments", *Nuclear Data Sheets*, Vol. 123, pp. 178-184 (2015), DOI:10.1016/j.nds.2014.12.031.

de Saint Jean, C., P. Archier, E. Privas and G. Noguère (2017), "On the use of Bayesian Monte-Carlo in evaluation of nuclear data", *EPJ Web of Conferences,* Vol. 146, 02007, DOI:10.1051/epjconf/201714602007.

de Saint Jean, C., P. Archier, E. Privas, G. Noguère, B. Habert and P. Tamagno (2018), "Evaluation of neutron-induced cross sections and their related covariances with physical constraints", *Nuclear Data Sheets*, Vol. 148, 383, DOI:10.1016/j.nds.2018.02.010.

Fiorito, L., G. Žerovnik, A. Stankovskiy, G. Van den Eynde and P.-E. Labeau (2017), "Nuclear data uncertainty propagation to integral responses using SANDY", *Annals of Nuclear Energy*, Vol. 101, pp. 359–366, DOI:10.1016/j.anucene.2016.11.026.

Fiorito, L., J. Dyrda and M. Fleming (2019), "JEFF-3.3 covariance application to ICSBEP using SANDY and NDAST", *European Physical Journal Web of Conferences*, Vol. 211, 07003, DOI:10.1051/epjconf/201921107003.

Fröhner, F.H. (2003), "Evaluation of data with systematic errors", *Nuclear Science & Engineering*, Vol. 145-3, pp. 342-353, DOI:10.13182/NSE03-A2387.

Habert, B., C. De Saint Jean, G. Noguère, L.l Lean and Y. Rugama (2010), "Retroactive generation of covariance matrix of nuclear model parameters using marginalization techniques", *Nuclear Science & Engineering*, Vol. 166, 276-287, DOI:10.13182/NSE09-85.

Helgesson, P., D. Rochman, H. Sjöstrand, E. Alhassan and A. Koning (2014), "UO_2 versus MOX: Propagated nuclear data uncertainty for k_{eff}, with burnup", *Nuclear Science and Engineering*, Vol. 177, 3, pp. 321–336, DOI:10.13182/NSE13-48.

Helgesson P., H. Sjöstrand, A. Koning, J. Rydén, D. Rochman, E. Alhassan and S. Pomp (2017a), "Combining Total Monte Carlo and Unified Monte Carlo: Bayesian nuclear data uncertainty quantification from auto-generated experimental covariances", *Progress in Nuclear Energy*, Vol. 96, pp. 76–96, DOI:10.1016/j.pnucene.2016.11.006.

Helgesson, P., H. Sjöstrand and D. Rochman (2017b), "Uncertainty-driven nuclear data evaluation including thermal (n,α) applied to ^{59}Ni", *Nuclear Data Sheets*, Bol. 145, pp. 1-24, DOI:10.1016/j.nds.2017.09.001.

Helgesson, P., D. Neudecker, H. Sjöstrand, M. Grosskopf, D. Smith and R. Capote (2018), "Assessment of novel techniques for nuclear data evaluation", Reactor Dosimetry: *16th International Symposium*, West Conshohocken, PA, pp. 105-116, DOI:10.1520/STP160820170087.

Hill, I., J. Gulliford, N. Soppera and M. Bossant (2014), "DICE 2013: New capabilities and data", *Proceedings of PHYSOR 2014* – The Role of Reactor Physics Toward a Sustainable Future, Kyoto, Japan, 28 September – 3 October 2014.

Ivanova, T., M. Nikolaev, K. Raskach, E. Rozhikhin and A. Tsiboulia (2003), "Analysis of the entire set of the HEU-SOL type experiments from the ICSBEP Handbook for correlation and systematic error", *Proceedings of the International Conference on Nuclear Criticality 2003*, Mito, Japan, 20-24 October 2003, Vol. 1, pp. 327-333.

Jeong, S., I. Hill, H. Kikusato and D. Lee (2017), "Creation of a database of uncertainties for ICSBEP Handbook and tool for covariance generation", *International Conference on Mathematics & Computational Methods Applied to Nuclear Science & Engineering*, Jeju, Korea, April 16-20, 2017.

Jouanne, C. (2014), "Sensitivity of the shielding benchmarks on variance-covariance data for scattering angular distributions", *Nuclear Data Sheets*, Vol. 118, pp. 384-387, DOI:10.1016/j.nds.2014.04.087.

Kodeli, I. and S. Slavič (2017), "SUSD3D computer code as part of the XSUN-2017 Windows interface environment for deterministic radiation transport and cross-section sensitivity-uncertainty analysis", *Science and Technology of Nuclear Installations*, Article 1264736, DOI:10.1155/2017/1264736.

Kodeli, I. (2018), "Comments on the status of modern covariance data based on different fission and fusion reactor studies", *European Physical Journal* N, Vol. 4, 46, DOI:10.1051/epjn/2018027.

Koning, A.J. and D. Rochman (2008), "Towards sustainable nuclear energy: Putting nuclear physics to work", *Annals of Nuclear Energy*, Vol. 35, pp. 2024-2030, DOI:10.1016/j.anucene.2008.06.004.

Koning, A.J. (2015), "Bayesian Monte Carlo method for nuclear data evaluation", *European Physical Journal* A, Vol. 51, 12, pp. 184, DOI:10.1140/epja/i2015-15184-x.

Koning, A.J., D. Rochman, J.-C. Sublet, N. Dzysiuk, M. Fleming and S. van der Marck (2019), "TENDL: Complete nuclear data library for innovative nuclear science and technology", *Nuclear Data Sheets*, Vol. 155, pp. 1–55, DOI:10.1016/j.nds.2019.01.002.

Kos, B., I. Kodeli, G. Žerovnik and L. Fiorito (2018), "Nuclear data uncertainty analysis of the TIARA iron benchmark using the random sampling code SANDY accelerated by the ADVANTG hybrid transport code", *Proceedings of PHYSOR 2018* – Reactor Physics Paving the Way Towards More Efficient Systems, Cancun, Mexico, 22-24 April 2018.

Lahaye, S. (2018), "Choice of positive distribution law for nuclear data", *EPJ Nuclear Sciences & Technologies*, 4. 38, DOI: 10.1051/epjn/2018047.

Larson, N. (2008), Updated Users Guide for SAMMY Multilevel R-matrix Fits to Neutron Data Using Bayes Equation, ORNL/TM-9179/R8, Oak Ridge National Laboratory, DOI:10.2172/941054.

Leeb, H., D. Neudecker and T. Srdinko (2008), "Consistent procedure for nuclear data evaluation based on modeling", *Nuclear Data Sheets,* Vol. 109, pp. 2762-2767, DOI:10.1016/j.nds.2008.11.006.

Litaize, O. and O. Serot (2009), "Investigation of phenomenological models for the Monte Carlo simulation of the prompt fission neutron and γ emission", *Physical Review* C 80, 044611, DOI:10.1103/PhysRevC.82.054616.

Lovell, A.E., P. Talou, I. Stetcu and K. Kelly (2020), "Correlations between fission fragment and neutron anisotropies in neutron-induced fission", *Physical Review* C, Vol. 102, 024621, DOI:10.1103/PhysRevC.102.024621.

MacFarlane, R.E., D. Muir, R. Boicourt, A. Kahler and J. Conlin (2019), The NJOY Nuclear Data Processing System, Version 2016, Los Alamos National Laboratory Report LA-UR-17-20093, DOI:10.2172/1338791.

Moxon, M.C., T. Ware and C.J. Dean (2010), REFIT-2009 A Least-Square Fitting Program for Resonance Analysis of Neutron Transmission, Capture, Fission and Scattering Data Users Guide for REFIT-2009-10, Report UK Nuclear Science Forum, UKNSF.

NEA (2014), Co-ordinated Evaluation of Plutonium-239 in the Resonance Region, NEA/NSC/WPEC/DOC(2014)447, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19478.

NEA (2013), "Methods and Issues for the Combined Use of Integral Experiments and Covariance Data", NEA/NSC/WPEC/DOC(2013)445, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19337.

NEA (2017), "Statistical Verification and Validation of the EXFOR Database: (n,γ), (n,n'), (n,2n), (n,p), (n,α) and Other Neutron-Induced Reaction Cross-Sections", NEA/DB/DOC(2017)1, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_19786.

NEA (2019a), International Criticality Safety Benchmark Evaluation Project Handbook (database), OECD Publishing, Paris, DOI: 10.1787/110ba6fc-en.

NEA (2019b), *International Co-operation in Nuclear Data Evaluation: An Extended Summary of the Collaborative International Evaluated Library Organisation (CIELO) Pilot Project*, NEA No. 7498, OECD Publishing, Paris, DOI:10.1016/j.nds.2018.02.003.

NEA (2019c), International Reactor Physics Evaluation Project Handbook (database), OECD Publishing, Paris, DOI: 10.1787/8d549c0f-en.

NEA (2020a), *Thermal Scattering Law S(α,β): Measurement, Evaluation and Application - International Evaluation Co-operation Volume 42*, NEA No. 7511, OECD Publishing, Paris, www.oecd-nea.org/jcms/pl_23901.

NEA (2020b), *Specifications for the Generalised Nuclear Database Structure (GNDS) Version 1.9,* NEA No. 7519, ISBN 978-92-6490-197-1, OECD Publishing, Paris, DOI:10.1787/94d5e451-en.

Neudecker, D., R. Frühwirth and H. Leeb (2012), "Peelle's pertinent puzzle: A fake due to improper analysis", *Nuclear Science & Engineering*, Vol. 170, pp. 54–60, DOI:10.13182/NSE11-20.

Neudecker, D. et al. (2018), "Evaluations of energy spectra of neutrons emitted promptly in neutron-induced fission of ^{235}U and ^{239}Pu", *Nuclear Data Sheets*, Vol. 148, pp. 1-142, DOI:10.1016/j.nds.2018.02.006.

Neudecker, D., M. White, D. Vaughan and G. Srinivasan (2020), "Validating nuclear data uncertainties obtained from a statistical analysis of experimental data with the physical uncertainty bounds method", *European Physical Journal* N, Vol. 6, 19, DOI:10.1051/epjn/2020007.

Wieselquist, W.A., R.A. Lefebvre and M.A. Jessee (eds.) (2020), SCALE Code System, Report ORNL/TM-2005/39, Version 6.2.4, Oak Ridge National Laboratory.

Palmiotti, G., M. Salvatores, H. Hiruta, M. Herman, P. Oblozinsky and M. Pigni (2011), "Use of covariance matrices in a consistent (multiscale) data assimilation for improvement of basic nuclear parameters in nuclear reactor applications: From meters to femtometers", *Journal of Korean Physics Society*, Vol. 59, 1123.

Rochman, D., A. Koning and S. van der Marck (2009), "Uncertainties for criticality-safety benchmarks and keff distributions", *Annals of Nuclear Energy*, Vol. 36, pp. 810-831, DOI:10.1016/j.anucene.2009.01.018.

Rochman, D., A. Koning, S. van der Marck, A. Hogenbirk and C. Sciolla (2011), "Nuclear data uncertainty propagation: Perturbation vs. Monte Carlo", *Annals of Nuclear Energy*, Vol. 38, 5, pp. 942–952, DOI:10.1016/j.anucene.2011.01.026.

Rochman, D. and A.J. Koning (2012), "Random adjustment of the H in H_2O neutron thermal scattering data", *Nuclear Science and Engineering*, Vol. 172, 3, pp. 287-299, DOI:10.13182/NSE12-2.

Rochman, D., W. Zwermann, S. van der Marck, A. Koning, H. Sjöstrand, P. Helgesson and B. Krzykacz-Hausmann (2014), "Efficient use of Monte Carlo: Uncertainty propagation", *Nuclear Science and Engineering*, Vol. 177, 3, pp. 337-349, DOI:10.13182/NSE13-32.

Rochman, D., E. Bauge, A. Vasiliev and H. Ferroukhi (2017), "Correlation nu-sigma-chi in the fast neutron range via integral information", *European Physical Journal* N, Vol. 3, pp. 14, DOI:10.1051/epjn/2017009.

Rochman, D., E. Bauge, A. Vasiliev, H. Ferroukhi, S. Pelloni, A. Koning and J.-C. Sublet (2018a), "Monte Carlo nuclear data adjustment via integral information", *European Physical Journal Plus*, Vol. 133, pp. 537, DOI:10.1140/epjp/i2018-12361-x.

Rochman, D., E. Bauge, A. Vasiliev, H. Herroukhi and G. Perret (2018b), "Nuclear data correlation between different isotopes via integral information", *European Physical Journal* N, Vol. 4, pp. 7, DOI:10.1051/epjn/2018006.

Rochman, D., A. Vasiliev, H. Ferroukhi, S. Pelloni, E. Bauge and A. Koning (2019), "Correlation nu-sigma for U-Pu in the thermal and resonance neutron range via integral information", *European Physical Journal Plus*, Vol. 133, pp. 453 DOI:10.1140/epjp/i2019-12875-7.

Schnabel, G., H. Sjöstrand, J. Hansson, D. Rochman, A. Koning and R. Capote (2021), "Conception and Software Implementation of a Nuclear Data Evaluation Pipeline," *Nuclear Data Sheets*, Vol. 173, pp. 239–284, DOI:10.1016/j.nds.2021.04.007.

Schnabel, G. (2017), "Fitting and analysis technique for inconsistent nuclear data", *Proceedings of the International Conference on Mathematics and Computational Methods applied to Nuclear Science and Engineering*, Jeju, Korea.

Schnabel, G. (2015), Large Scale Bayesian Nuclear Data Evaluation with Consistent Model Defects, Doctoral Thesis, Vienna University of Technology, AC12392895.

Shibata, K. et al. (2011), "JENDL-4.0: A new library for nuclear science and engineering", *Journal of Nuclear Science and Technology*, Vol. 48, pp. 1-30, DOI:10.1080/18811248.2011.9711675.

Siefman, D. (2019) Development and Application of Data Assimilation Methods in Reactor Physics, Doctoral Thesis, École Polytechnique Fédérale de Lausanne, Switzerland, DOI:10.5075/epfl-thesis-7525.

Sjöstrand, H. and G. Schnabel (2019), "Monte Carlo integral adjustment of nuclear data libraries – experimental covariances and inconsistent data", *European Physical Journal Web of Conferences*, Vol. 211, 07007, DOI:10.1051/epjconf/201921107007.

Sobes, V., C. de Saint Jean, D. Rochman, O. Cabellos, A. Holcomb, E. Bauge, R. Capote, A Trkov, and M. Fleming (2021), "WPEC Subgroup 44 Computational Inter-comparison Exercise on Correlations in Nuclear Data Libraries", *Annals of Nuclear Energy,* Vol. 164, 108605, DOI: 10.1016/j.anucene.2021.108605.

Sublet, J. et al (2019), "Neutron-induced damage simulations: Beyond defect production cross-section, displacement per atom and iron-based metrics", *European Physical Journal Plus*, Vol. 134, pp. 350, DOI:10.1140/epjp/i2019-12758-y.

Vaughan, D.E. and D.L. Preston (2015), Physical Uncertainty Bounds (PUB), Los Alamos National Laboratory Report LA-UR-14-20441, DOI:10.2172/1209314.

Vogt, R., J. Randrup, J. Pruet and W. Younes (2009), "Event-by-event study of prompt neutrons from ^{239}Pu(n,f)", *Physical Review* C 80, 044611, DOI:10.1103/PhysRevC.80.044611.

Zerkin, V. (2012), "Web tool for constructing a covariance matrix from EXFOR uncertainties", *European Physical Journal Web of Conferences*, Vol. 27, 9, DOI:10.1051/epjconf/20122700009.

Žerovnik, G., A. Trkov, I. Kodeli, R. Capote and D. Smith (2015), "Random sampling of correlated parameters – A consistent solution for unfavourable conditions", *Nuclear Data Sheets*, Vol. 123, pp.185-190, Academic Press, Cambridge, MA, DOI:10.1016/j.nds.2014.12.032.