Nuclear Reactors, Evaluations, Library Development

Patrick Talou
T-2, Nuclear Physics Group, Theoretical Division
Los Alamos National Laboratory, USA
talou@lanl.gov

Abstract:
This lecture reviews what constitutes a nuclear data evaluation, how evaluated data files come to life, and why they remain so important to this day. After discussing what exactly is a nuclear data evaluation, we introduce the nuclear data evaluation process itself, in which state-of-the-art experimental data and theoretical calculations are combined to provide a snapshot of our current understanding of the nuclear reactions of interest. This knowledge is finally encapsulated into a computer file that can be used by transport simulation codes for applications. We then describe in more details how experimental data are collected, analyzed and combined statistically. The role of theoretical models of nuclear reactions is then discussed through two examples: nuclear fission cross sections, and prompt fission neutron spectra. The quantification of uncertainties stemming from both experimental data and model calculations plays a significant role in the final evaluated results. We discuss how uncertain nuclear data are combined, in particular through a Bayesian inference scheme. Finally, we briefly mention the formatting and processing of nuclear data files, and how they are used in transport codes for application needs. We conclude by providing a glimpse on what can be expected in the next 5-10 years in the “Art and Science” of nuclear data evaluation.

1 Introduction

Nuclear applications, ranging from basic nuclear physics experiments, nuclear reactors, nuclear waste management, nuclear medicine, nucleosynthesis, non-proliferation, space exploration, etc., all rely on advanced computer simulations that make significant use of evaluated nuclear data. With the advent of very large modern computing power, those simulations are increasingly complex and detailed. As an example, the CASL initiative [1] in the United States is a Department of Energy (DoE) “Energy Innovation Hub” established to perform advanced modeling and simulation for commercial nuclear reactors. It uses High Performance Computing capabilities to perform very detailed 3D simulations (see Fig. 1) of neutronics, hydraulic, and more generally multi-physics problems related to the operation of a light water reactor.

The design of Generation-IV reactors as well as small and modular reactors (SMR) with innovative fuels necessitates a fresh look at some of the evaluated nuclear data of lesser importance in the past, e.g., for minor actinides. The efficient and safe development of the Th-U fuel cycle also relies on nuclear data that did not draw much attention until recently. In the realm of non-proliferation and treaty verification, nuclear data related to the fission
The Consortium for Advanced Simulations of Light-water reactors (CASL) [1] is interested in performing large-scale numerical simulations of light-water reactors using the unparalleled power of modern super-computers. This figure displays the result of MPACT simulations of the fission rate distribution in fuel rods for the initial criticality of the reactor at zero power.

process appear to be more and more relevant for advanced simulations and the development of new signatures. Nuclear medicine is a fast growing field both in terms of imaging and radiotherapy. In both cases, accurate nuclear data are required to provide accurate simulations that complement and guide the medical procedures.

In all those nuclear applications, accurate and predictive simulations require precise and complete nuclear reaction data that are used to follow the interaction and transport of particles through matter. Such evaluated data are made available through nuclear data libraries, such as the U.S. ENDF/B-VII.1 library [2], the OECD/European JEFF-3.1.1 library [3], or the Japanese JENDL-4.0 library [4], to name but a few.

How are these evaluated data generated is the purpose of this course.

A nuclear data evaluation is often the result of an artful combination of experimental differential information, theoretical insight, model predictions, integral benchmarks as well as the “expertise” of the evaluator. In this lecture, I will describe how differential experimental data are analyzed through statistical techniques and stress the particularly important role of systematic, often unrecognized experimental errors. Proper account for these uncertainties can dramatically alter the evaluated covariance matrices that are part of the evaluated libraries.

Bayesian statistical inference is ubiquitous in many scientific, technological, economic, sociological studies, and is also commonly used in nuclear data evaluation works. It is used mostly to combine knowledge acquired from several sources, such as experimental data and model calculations. Bayesian updating schemes can be used to constrain the mean values and uncertainties of model input parameters, which in turn can be used to predict physical quantities where no experimental data exist.

Nuclear reaction models are often used to predict reaction cross sections, angular and energy dependent secondary particles, multiplicity, production rates, etc. At sufficiently high excitation energies, the statistical Hauser-Feshbach model [5], supplemented by pre-equilibrium effects and width fluctuation corrections, is commonly used in codes such as GNASH [6], EM-
Figure 2: The U-235 (n,2n) and U-238 (n,tot) evaluated cross sections, and their one-sigma error bands, are compared to experimental data.

PIRE [7], TALYS [8], CoH [9], CCONE [10]. Many phenomenological models are present in these codes, and their parameters often tuned to appropriate experimental data. Examples of physics models commonly used are briefly discussed in Section 2.2.

Figure 2 depicts two examples of important and well-studied cross sections with their evaluated mean values and one-sigma uncertainty bands, in comparison to available experimental data. As one can see in the case of $^{235}$U (n,2n) cross section, experimental data may exit but only in a limited energy range, and have to be complemented with model predictions.

To be of broad and general use, an evaluated nuclear data file has to be as complete as possible. Transport codes, such as the well-known Monte Carlo MCNP code [11], use processed formatted evaluated libraries, and need physical processes to be fully described for a wide range of target nuclei, projectiles and reaction energies. General purpose libraries are relatively complete for incident neutrons, for example. Completeness does not equate quality however, and the quality of a particular evaluated file often reflects the importance of this particular isotope for specific nuclear applications.

In this lecture, I will review the nuclear data evaluation process, as briefly described above, and discuss several important applications, including Gen-IV type reactors, small and modular reactors, accelerator-driven systems, non-proliferation, to illustrate the importance of evaluated nuclear data libraries. I will conclude by discussing advances in physics models, experimental data, nuclear reaction computer simulations, nuclear data formats, etc. I will also briefly mention the recent International evaluation effort CIELO, which represents an attempt at pulling together worldwide evaluation resources to establish a state-of-the-art evaluated nuclear data library.

2 How to Generate a Nuclear Data Evaluation

Note: the evaluation of nuclear data is a vast topic that spans nuclear reactions, nuclear structure, nuclear masses, decay data, product yields, medical radiation doses, astrophysics rates, etc. While being specific, the evaluation of each type of data still follows a set of similar rules and methods. In this paper, we focus primarily on the evaluation of nuclear
reaction data.
Nuclear data evaluations should reflect our state of knowledge on particular reactions on a
given target nucleus. In some cases, experimental data inform us. In other cases, theory and
modeling alone can provide some physical constraints without the need for new measure-
ments. In most cases, both theory and experiment play a role in the elaboration of a nuclear
data evaluation. While details of an evaluation differ for different nuclei- from light to heavy,
and for different incident particle energies- from epithermal to GeV, the same general ideas
apply, and will be described below.

2.1 Analyzing Experimental Data

One of the first steps in performing a nuclear data evaluation is to collect and analyze all
experimental data available regarding a particular isotope and reaction of interest. Often,
a new evaluation really means a reevaluation of an existing file, corresponding to a specific
isotope, and including several reactions of interest, their products, spectrum, angular distri-
butions, etc. However, they are often limited in incident energy and quantities of interest.
For instance, We may be interested in the neutron-induced fission reaction on $^{241}$Am from
thermal to 20 MeV incident neutron energies only. In this case, only this particular reaction
as well as all competing reaction channels will be of interest. At low energies, the capture
reaction will be the only competing channel, and will mostly vanish with increasing incident
energies. At higher energies, neutron inelastic channels will be the first ones to compete,
while other channels, e.g., proton or alpha emissions, may open up at even higher energies.

Databases of Experimental Data and Retrieval Tools

An important tool at the disposal of the nuclear data evaluator is the existence of databases
of experimental nuclear data. In the past, this information used to be stored in books
(CINDA, ...). Nowadays, computer relational databases have replaced advantageously those
paper-based data, providing advanced search-engines to quickly and efficiently retrieve a set
of experimental data corresponding to a set of well-defined search criteria.

Nuclear data centers spread across the world are in charge of collecting new experimental
data, fixing existing databases, and retrieving old experimental data that have not been saved
digitally so far. Nuclear data centers include: the National Nuclear Data Center (NNDC) at
Brookhaven National Laboratory [12] in the USA; the Nuclear Energy Agency [13], located
in Paris, France; the Nuclear Data Services at the International Atomic Energy Agency [14]
in Vienna, Austria, etc. Those centers are crucial for the dissemination of nuclear data, and
are used by the entire nuclear physics community on a regular basis.

Modern relational databases can be used to quickly search for particular data sets cor-
responding to a given nuclear reaction of interest. Their use is made easy through web
interfaces that are fairly straightforward to use. Figure 3 provides a screenshot of such a
web interface at the NNDC website. Experimental data sets are stored using the EXFOR
format, which defines keywords and sections for easy retrieval of information. Unfortunately,
the rules for writing in this format are rather loose, which complicates the process of accu-
rately extracting specific data in a specific format. Nevertheless, those tools remain crucial
to the evaluator’s toolkit.
Figure 3: A screenshot of the request form to access the experimental nuclear reaction database (EXFOR), available through the website (http://www.nndc.bnl.gov) at the National Nuclear Data Center (NNDC) at Brookhaven National Laboratory. Relational databases have become an important tool for the evaluator to quickly access a large amount of experimental information on nuclear reaction data.

In addition to retrieval forms, nuclear data centers often also provide plotting capabilities (e.g., Sigma at NNDC, JANIS at NEA) as well as relatively simple physics checking codes that allow for a quick and now relatively automatic way for checking a newly submitted evaluation.

Dealing with Old and Discrepant Data Sets

In an ideal situation, all data sets relative to a given physical quantity of interest would report the same value, within their quoted error bars. In other words, those data sets would be consistent. Unfortunately, this situation is not as common as one could expect naively. In fact, experimental measurements are often impacted by known as well as unknown uncertainties, errors and correction factors. A “blind” use of experimental data extracted from the EXFOR database is likely to lead to erroneous answers, and much caution should be placed on carefully reading the information either stored in the EXFOR files, or/and in the original publications and reports from the authors of the experiments themselves.

Another reason to be cautious when collecting and analyzing experimental data sets lies in the fact that the reported data are actually not necessarily what has been measured. As an example, most cross-section measurements are actually measurements of the ratio of two cross-sections. There is a very good reason for measuring this type of data. In a typical cross-section measurement, important sources of uncertainties, such as the absolute neutron flux in a neutron-induced reaction, simply cancel out when ratio data are considered. Such measurements almost always use a very well-known cross-section for the denominator, so that the ratio data can be easily transformed into cross-sections. The final uncertainties will be a convolution of the ratio uncertainties with the supposedly small uncertainties of the cross-sections in the denominator. A subset of cross-sections that qualify for this criterium constitutes the so-called “standard” cross-sections, which are being discussed in Section 2.4.
To illustrate this point, let us consider the ratio of the neutron-induced fission cross-sections of $^{239}$Pu over $^{235}$U in the keV to 20 MeV region. In this energy region, the $^{235}$U (n,f) cross-section is a standard, and has been reported with accuracy better than 1% for most incident energies. Now, an older measurement may report experimental data for the $^{239}$Pu (n,f) cross-section, albeit the data taken were really ratio data. In such a case, the ratio raw data were transformed by the authors to cross-section data using an evaluated cross-section for $^{235}$U (n,f). However, chances are that this cross-section is different from what is now considered the standard. To get back to the correct experimental data, the evaluator would have to rely on the original ratio data, if available, and not on the inferred cross-section data.

![Pu-239/U-235 (n,f) Cross Sections Ratio](image)

**Figure 4:** Ratio of $^{239}$Pu and $^{235}$U neutron-induced fission cross sections. By performing ratio measurements, experimentalists bypass the difficult problem of accurately assessing the neutron flux, which is often a dominant source of uncertainties in cross section measurements. The reaction $^{235}$U(n,f) is a “standard” (see Section 2.4) that has often been used in such ratio measurements. In this _atypical_ example, the ratio of Pu239 to U235 neutron-induced fission cross sections has been measured very accurately and consistently in various experiments.

Estimating realistic uncertainties associated with the reported data, beyond those sometimes reported by the authors themselves, is an even much more challenging task, which can only be addressed subjectively. We discuss this question in more details in the next section.

Another very problematic issue faced by any evaluator concerns the lack of proper documentation of mostly older, but some recent ones as well, experiments. The less information is given, the more subjective the interpretation of the evaluator becomes. When other, more recent and better documented experiments exist, it often leads to the simple rejection or, more elegantly, the weighing out of the older experiment by increasing significantly the reported uncertainties. When no other experiment exists, the decision becomes more difficult... and subjective. It becomes even more important for the evaluator to carefully review the measured data and any indirect documentation that can be found.
A difficult and unfortunately common problem arises when two or more data sets are discrepant within one to two standard deviations, without a clear indication of where a problem could have occurred in (at least) one of the experiments. Sometimes the solution has to be found elsewhere, where indirect information on this particular reaction plays a role. Of course, it is always possible to increase the error bars of each experimental point, naturally leading to a decrease of the $\chi^2$/d.o.f., hence to a seemingly better agreement. However, this solution is rarely the correct one, as systematic unknown errors, which tend to be the source of discrepancies, would bring correlations among the measured points, which would not be captured by the standard deviations alone.

As an example, Fig. 5 shows the evaluated (ENDF/B-VII.0) neutron-induced fission reaction on $^{243}$Am in comparison with experimental data sets. Significant discrepancies exist between various data sets, and a systematic shift is observed in the 1—20 MeV incident energy range. As we will be discussed later, theoretical models of fission cross sections often remain qualitative and input parameters are tuned to available experimental data, so they are of little help when dealing with discrepant fission cross sections. In this particular case, the use of integral data (see Section 3) helped determine the correct fission cross section normalization, corresponding to the ENDF/B-VII values here.

![Figure 5: Evaluated (ENDF/B-VII.0) and experimental data on the neutron-induced fission cross sections of $^{243}$Am. In the 1-10 MeV incident neutron energy range, the experimental data sets split into two main groups—the surrogate reaction results, also shown here, are expected to be less accurate in this case and fall outside our discussion. Those two groups appear to be mostly shifted by a rather constant factor across this energy range. Integral data were used to infer the correct magnitude for this cross section. Details in the text.](image-url)
Quantifying Experimental Uncertainties

This is perhaps one of the most difficult and time-consuming tasks to be performed by an evaluator. It is common practice for experimentalists to present their results in conferences and conference proceedings, and in peer-reviewed journals. Conference proceedings often deal with preliminary data and can rarely be considered as usable data in their published form. It is expected that at some point, the author(s) submit their experimental results to the scrutiny of peers in a refereed journal. At this point, they would (should) also send their final data points to the EXFOR database compilers. However, the documentation of the exact experimental setup and procedure is often lacking, in part due to the reluctance of physics journals to publish long and detailed explanations. Some additional explanations can be provided in the EXFOR database, but they are often very limited, if given at all. The job of an evaluator therefore becomes the one of an investigator trying to find clues to why discrepancies are observed or reported uncertainties are so small.

Experimental uncertainties often fall into two categories, although the reality is a bit more complex. Statistical uncertainties constitute the first category, and decrease with the number of counts. On the other hand, systematic uncertainties do not disappear when the number of counts is increased. Those are due for instance to detector efficiency, impurities in the sample, etc. While statistical uncertainties are commonly reported in the published literature, it is often more difficult to find comprehensive information on the second type of uncertainties.

As an example, the different sources of (known) uncertainties impacting Knitter’s measurement [15] of the prompt fission neutron spectrum (PFNS) of $^{239}$Pu (n,f) induced by 0.215 MeV neutrons are shown in Fig. 6, as well as the resulting correlation matrix. All uncertainties, except for the statistical ones, contribute to the non-zero off-diagonal elements in the correlation matrix. Correctly assessing those correlations is very important to get realistic
final evaluated uncertainties. In this particular case, neglecting them leads to very small evaluated uncertainties on the PFNS.

2.2 Nuclear Reaction Model Calculations

An evaluation is only very rarely based entirely on experimental data sets for at least one major reason: they only cover part of the energies, the reactions or the isotopes of interest. As mentioned in the introduction, a transport code such as MCNP [11] requires a complete set of reaction probabilities and other data to provide reasonable answers. For instance, if reactions such as \((n,f)\) or \((n,\gamma)\) are evaluated over a limited energy range only, what happens if a neutron is produced outside this energy range? The code would crash or would ignore the reaction for which no data is available. Both situations are very problematic.

Experiments cannot measure all required nuclear data at all energies. It would certainly be prohibitively expensive, as well as impossible at times- for instance, some isotopic targets simply cannot be made, in particular if they decay very quickly, e.g., isomeric states. Therefore, evaluated nuclear data files have to rely on other sources of information. Nuclear theories and models fill this important gap.

A complete description of the nuclear reaction theories used in the nuclear data evaluation process is outside the scope of this lecture. It would also be redundant with the lectures given during the school by S. Hilaire [16] on nuclear reaction theories in the fast energy range as used in the TALYS code [8], and by F. Gunsing [17] on the R-matrix theory used for nuclear reactions in light nuclei and in the analysis and fitting of resonances in the medium to heavy nuclei.

Instead, and for illustrative purposes, I will limit the discussion to two nuclear reaction modeling topics: (1) fission cross sections; and (2) prompt fission neutron spectrum (PFNS). Both are very relevant to several important ongoing research programs, through theory and experiment alike. Also, the PFNS has been used in the practical mini-workshop sessions.

Nuclear Fission Cross Sections

Discovered more than 75 years ago, the nuclear fission process remains a formidable challenge for theorists and experimentalists alike. While many qualitative features of fission are reasonably well understood, a complete, correlated and quantitative description remains elusive. Nuclear fission is a very rich and complex process, which can only be fully understood and quantified through quantum mechanics. The fission cross sections represent the probability that a particular reaction (incident particle, energy, polarization, etc.) on a heavy nucleus would lead to the fission or break-up of this nucleus into two or more fragments. Here, we restrict ourselves to the study of low-energy nuclear fission, in which the most dominant process leads to the formation of two complementary fragments. Note that ternary fission in which an alpha particle is emitted accounts for no more than 0.5% of the total fission events.

Only a few months following the experimental discovery of the fission process, Bohr and Wheeler [18] published their seminal paper on “The mechanism of nuclear fission”, in which they discussed the theory of fission cross sections based on the liquid drop model and the number of transition levels available on top of the barrier. By analogy with a liquid drop that
can split into two smaller drops, a heavy nucleus can undergo such a transformation. If one represents the energy of the system as a function of its shape and in particular its degree of deformation on its path to fission, a potential energy barrier appears between the system in its original, near spherical ground-state configuration and the two fragments configuration. This is shown in their original drawing and reproduced in Fig. 7.

Over the years, many refinements have been made to this rather simple picture, but the general framework remains. In modern versions of the Finite-Range Liquid-Drop Model (FRLDM) [19], the shape of the nucleus is described using 5 parameters (see Fig. 8 left), leading to complicated fission paths in this five-dimensional landscape. Modern simulations have shown that indeed 5 parameters are a minimum to properly describe the fission paths and infer the fission barrier characteristics. Large-scale microscopic calculations using the Hartree-Fock formalism (see Fig. 8 right) are also being performed [20, 21, 22] to estimate fission paths, fission barriers, fission fragment yields, etc. The presence of an isomeric well or class-II states on the path to fission complicates the modeling of fission cross sections significantly.

In the formal R-matrix theory of fission cross sections [23, 24], calculating fission probabilities requires the knowledge of many quantities: fission barrier heights and widths, fission transition states at the saddle points, characteristics of the second minimum, inertia tensor, neutron and proton pairing gaps as a function of nuclear deformation, coupling strengths between the two wells, etc. If more than one fission paths or modes contribute, then the picture is even more complicated.

It is therefore no surprise that many model input parameters enter in the calculation of the fission cross sections. While physical constraints and experimental data do exist on many of these parameters, a wide selection of different parameter sets can provide similar results. In other words, there is no unique solution of parameter set to describe those cross sections. By
Figure 8: (a) Nuclear shape parameterization used in the finite-range liquid-drop model [19], and (b) potential energy surface in the fission of $^{258}\text{Fm}$ calculated in the Hartree-Fock model using the Skyrme nucleon-nucleon interaction [20].

describing more than just cross sections, new models would be more selective and eliminate some spurious solutions.

**Prompt Fission Neutron Spectrum**

As another example, the average prompt fission neutron spectrum represents the average distribution in energy of the prompt neutrons emitted right after scission from excited primary fission fragments. Those neutrons are emitted within $10^{-14}$ s of the scission time. Simplified models have existed for a long time. To this day, the most widely used model for evaluation purposes remains the so-called Los Alamos or Madland-Nix model [25] or a more sophisticated version of it. In this model, the Weisskopf’s evaporation theory of nuclear reactions [26] is used to describe the emission of the neutrons in the center-of-mass of the fission fragments. A triangular distribution of effective nuclear temperatures in the fragments is assumed, closely resembling the distribution already observed by Terrell [27]. Assuming an isotropic emission from the fragments, those neutrons are kinematically boosted in the laboratory frame, and the average prompt fission neutron spectrum in the laboratory frame reads

$$N(E) = \frac{1}{2\sqrt{E_f T_m^2}} \int \frac{(\sqrt{E} + \sqrt{E_f})^2}{(\sqrt{E} - \sqrt{E_f})^2} \, d\epsilon \sigma_c(\epsilon) \sqrt{\epsilon} \int_0^{T_m} dT k(T) \exp(-\epsilon/T),$$  \hspace{1cm} (1)$$

where $E_f$ is the fragment kinetic energy per nucleon, $\sigma_c(\epsilon)$ is the energy-dependent cross section for the inverse compound formation process, $T_m$ is the maximum nuclear temperature in the fragments, and $k(T)$ is the triangular temperature distribution. In the original version of the Los Alamos model [25], the temperatures in the light and heavy fragments are set equal, and the total average neutron spectrum is just the average of the light and heavy spectra. Equation (1) was also computed for a few fragmentations only.
Figure 9 shows the average prompt fission neutron spectrum in the case of 0.5 MeV incident neutrons on $^{239}$Pu, on both an absolute scale (left figure) and in ratio to a Maxwellian at temperature $T=1.42$ MeV. In a first approximation, a PFNS resembles a Maxwellian function. It is therefore of common usage to represent a PFNS as a ratio to a Maxwellian in order to better identify differences between experimental data, model calculations, and evaluated results at both ends of the spectrum— at the lowest outgoing energies below 500 keV, where multiple scattering corrections dominate, and at the highest energies, above 10 MeV, where poor statistics become an issue.

While the Los Alamos model represents a convenient and efficient way to represent a PFNS with only a handful of parameters, the reality is of course much more complex: in a typical fission experiment, fragments are produced in a variety of mass and charge splits, a wide range of kinetic energies, and various initial configurations in excitation energy, spin and parity. In a simple picture, prompt neutrons are evaporated from the fully accelerated fragments, and gamma rays follow to get rid of residual rotational energy. Contributions from the emission of neutrons near the scission point and gamma rays emitted prior to the emission of neutrons should also be taken into account. At incident neutron energies above the neutron separation energy, second-chance ($n, n'f$) fission processes can occur in which a neutron is first emitted leading to a residual nucleus, which then fissions. Higher excitation energies lead to multi-chance fissions. Pre-equilibrium effects also become important above about 10 MeV incident energy, and lead to anisotropic neutron angular distributions.

Recent efforts [28, 29, 30, 31] have focused on Monte Carlo simulations of the fission fragment de-excitation process on an event-by-event basis. This powerful approach can provide distribution and correlation data on prompt fission neutrons and photons, inaccessible through the Los Alamos model. In addition, it provides natural correlations between quantities such as the average neutron multiplicity $\bar{\nu}$, multiplicity distribution $P(\nu)$, mass-dependent average multiplicity $\nu(A)$, etc., and the average neutron spectrum calculated above. It implies that
different experimental data can be used to constrain a given quantity such as the PFNS.

2.3 Quantifying Uncertainties from Model Calculations

Model predictions do come with uncertainties. After all, these are just models, and not constants of Nature. Uncertainties associated with model predictions come in two flavors:

[1] Uncertainty in the model input parameters;

[2] Limitations or deficiencies of the models.

Unfortunately, distinguishing those two types of uncertainties in real applications is often very difficult. Let us first consider the extreme case in which the model is assumed to be perfect and that uncertainties only originate in our limited knowledge of the model input parameters. Existing experimental data can be used to constrain the prior parameter space and infer the “best values” for the input parameters as well as a probability distribution function (PDF) around them. The simplest and most direct, but also the most time-consuming, approach is the use of brute force Monte Carlo sampling of the parameter space. Random sets of input parameters are then fed into the model calculations, which produce numerical results that can be compared to experimental data. The mean values and PDF are then found by minimizing the $\chi^2$ per degree-of-freedom, or any other measure of discrepancy between calculations and experiments.

Other approaches can be used to infer uncertainties stemming from model calculations. For instance, in a linear and first-order approximation, the sensitivity of the model predictions to a changes in parameter values can be used to quickly assess the impact of model parameter uncertainties on the final uncertainty of the predicted quantity. Relative sensitivity coefficients defined as

$$S_{ij} = \frac{p_j}{M_i} \frac{\partial M_i}{\partial p_j},$$

compute the sensitivity of the model value at point $i$ (e.g., energy $E_i$) to a change in the parameter $p_j$. Those coefficients are then folded with the covariance matrix $C$ of the model parameters through the so-called “sandwich” rule:

$$\delta = SCST$$

to provide the standard deviation on the integral quantity of interest. If the model is not perfect- which it never is, then a bias is introduced in the evaluated result. For instance, trying to fit experimental data that follow a Poisson distribution with a Gaussian function will lead to a bias in the final mean value and width of the Gaussian distribution.

However, this approach, which is sometimes referred to as “parametric model uncertainties”, is very common due in large part to the difficulty in assessing the imperfections of the model.
2.4 A Learning Process

A nuclear data evaluation should encapsulate all information relevant to particular physical quantities related to nuclear reactions on a given isotope. As we have seen, this information could come from either experimental or theoretical sources. In this Section, we will discuss several frameworks to combine those two sources, in particular the Bayesian statistical approach and various Monte Carlo techniques. The special role played by “standard” nuclear reactions will be discussed as well. Finally, the difficult question of model errors will be mentioned.

The Bayesian Inference Scheme

Bayes’ theorem [32] provides a simple and elegant mathematical tool to assimilate knowledge. Simply derived from conditional probability equations, it reads as

\[
P(H|D,I) = \frac{P(D|H,I) \times P(H|I)}{P(D|I)} \tag{4}
\]

where \(H\) represents an hypothesis that is being tested, \(D\) some data (experimental or model-calculated), and \(I\) some independent prior information about the hypothesis. The term \(P(H|I)\) represents the prior probability that the hypothesis \(H\) is true prior to acquiring the new data \(D\). The term \(P(D|H,I)\) is the likelihood function that determines the probability of observing the data \(D\) if the hypothesis \(H\) were true. The product of those two terms gives, except for a normalization factor, the posterior probability \(P(H|D,I)\) of \(H\) being true given the prior evidence \(I\) and the new data \(D\).

Applying Bayes to Nuclear Data Evaluations

In the context of nuclear data evaluation, Bayes’s theorem can be applied iteratively to assimilate nuclear data sets for the same (or a simple function of a) physical quantity. Note that in this case, incorporating data sets does not need to follow the chronological order in which the data were acquired in the first place.

The Kalman filter was developed by R.E. Kalman [33] and was first used in the Apollo space navigation computer. Applying the Kalman filter to the problem of tracking trajectories, and using a model for the trajectory of the moving probe, the filter would first predict its position at a given time based on the knowledge of a prior position and velocity, and then correct/modify the parameters of the model according to a new measurement of the probe position. In other words, the Kalman filter is a Predictor-Corrector method, assimilating new information as it becomes available.

In the context of nuclear data evaluations, there is no dynamics or time involved, and the equations simplify accordingly. Let us assume that we are evaluating a reaction cross-section that can be modeled using a set of parameters \(x\) by \(\sigma = f(x)\). Let us further assume that the model is linear and that to a first-order Taylor series expansion around the central parameter values \(x_0\), one can write:

\[
\sigma = f(x) + e \simeq f(x_0) + C(x - x_0) + e, \tag{5}
\]
where \( C \) is the sensitivity matrix and \( e \) is some small correction. The sensitivity matrix measures the change in the model-calculated values driven by a change in the parameter values. In this case, and following the notations of Kawano [34], one can write the following equations for the Kalman filter:

\[
\begin{align*}
x_1 &= x_0 + XC^T(CXC^T + V)^{-1}(y - f(x_0)) \\
P &= X - XC^T(CXC^T + V)^{-1}CX.
\end{align*}
\]

The terms \((x_0, X)\) represent the prior parameters and associated covariance matrix, and \((x_1, P)\) represent the posterior parameter values and covariance matrix. The experimental data and uncertainties are given by \((y, V)\). Such a simplified, linear and first-order Kalman filter has been implemented in the KALMAN code [35] and has been used extensively to compute cross-section covariance matrices [36] for the ENDF/B-VII.1 library. It also been applied more recently to the evaluation of prompt fission neutron spectra [37, 38]. More details were given in the applied mini-workshop accompanying these lecture notes.

![Figure 10: The neutron-induced fission cross section of Pu-238 (left) was evaluated through a least-square fit of all available experimental data sets. Its correlation matrix (right) is therefore close to being diagonal.](image)

For nuclear data evaluations, the Kalman filter can be used in several ways: (1) experimental data only; (2) model predictions only; (3) experimental data and model predictions. Because models rely on a relatively small number of parameters to compute large sets of data, the correlations in the final evaluated values tend to be stronger in the case (2) than in the case (1), while case (3) lies somewhere in between these two extreme cases. As an example, the Pu-238 neutron-induced fission cross section is shown in Fig. 10 along with its evaluated correlation matrix. In this case, the cross section was evaluated using experimental data only, leading to a mostly uncorrelated matrix. On the contrary, in the case of the neutron-capture cross section of \(^{235}\text{U}\) (see Fig. 11), strongly correlated model calculations dominate the evaluation, and the resulting correlation matrix therefore shows strong off-diagonal positive elements.

**Monte Carlo Techniques**
Figure 11: The neutron-induced capture cross section of U-235 was evaluated using the KALMAN code [35] and a combination of model predictions and experimental data. In this case, the evaluated correlation matrix is strongly correlated, as opposed to what is observed in Fig. 10.

In contrast to the deterministic Kalman filter approach discussed above, various uncertainty quantification techniques have been devised based on the Monte Carlo technique: “Filtered Monte Carlo”, “Backward-Forward Monte Carlo” [39], “Total Monte Carlo” [40], “Unified Monte Carlo” [41]. In all cases, the general idea is to randomly sample the phase space of the parameter values, to somehow measure the “distance” between the model predictions and the experimental data, and then to minimize those discrepancies to find the “best” model parameter values. Thanks to the dramatic increase in computing power, those techniques once prohibitive are becoming more and more amenable to large-scale calculations. They are also not restricted to approximations such as the first-order linear approximation made in the KALMAN code\textsuperscript{1}, and take account of non-linear response functions automatically. They are also not restricted to the implicit assumption in covariance matrices of Gaussian distributions of uncertainties.

The Total Monte Carlo (TMC) technique goes a step further, and bypasses the need for ENDF formatting the results of an evaluation to directly transport the calculated nuclear data into transport simulations. By varying the model input parameters, the TMC technique can be used to directly infer the distribution of simulated integral quantities, such as the multiplication factor $k_{\text{eff}}$ in a critical assembly. As shown in [42], the final distribution can be skewed and non-Gaussian, which is quite important in assessing the safety and economical margins of a reactor.

The “Standards”

Some nuclear data evaluations rely on experimental data only, on purpose. Those reactions are considered “standards” and have been most recently evaluated and reported by Carlson et al. [43]. Standard reaction cross sections are very useful since they are thought to be

\textsuperscript{1}A second-order approximation has be implemented in KALMAN for testing purposes only.
very well known, and therefore useful for calibration and ratio measurements. The incident neutron fluence is usually very difficult to measure precisely and therefore represents a sizable component of the total experimental uncertainty. By performing a calibration or a ratio measurement to a standard reaction, the need to measure the neutron fluence is eliminated. The current standard reaction cross sections include:

\[ H(n, n), \ ^{3}\text{He}(n, p), \ ^{6}\text{Li}(n, t), \ ^{10}\text{B}(n, \alpha), \ ^{10}\text{B}(n, \alpha\gamma), \ ^{10}\text{B}(n, \alpha\gamma), \text{C}(n, n), \text{Au}(n, \gamma), \text{and} \ ^{235}\text{U}(n, f). \]

Thermal constants (e.g., fission neutron multiplicities at thermal energy) are also part of the Standards. Other reaction cross sections, e.g., \(^{238}\text{U}(n, f)\), have been used in the generalized least-square fit, but are not considered standards. The prompt fission neutron spectrum of \(^{252}\text{Cf} \) spontaneous fission is also a standard, which is largely used to determine the energy efficiency of neutron detectors.

**Model Errors**

In all the techniques mentioned above, except in the special case of the “standard” reactions, theoretical models play a key role in the final evaluated result. An implicit assumption is that any uncertainty originating in the use of the model is due to uncertain model input parameters. In other words, it means that the model is considered to represent reality perfectly and that only uncertainties in the parameters prevent the model from accurately determining the physical quantity of interest with absolute accuracy. This is of course a false and oversimplified statement.

By blaming the model input parameter uncertainties for the total model uncertainty, we are making two important mistakes: the first one is that the uncertainties and in particular correlations are not properly evaluated and can bias the final mean values of the quantity of interest. The second one is that one might very well underestimate the true uncertainties by a sometimes dramatically large amount. This is particularly worrisome when model predictions are carried far from where their input parameters were constrained by experimental data. By ignoring or minimizing the effects of a missing physics process, one may very well end up with the wrong answer and small uncertainty. While some statistical techniques can be used to identify outliers and systematic deviations of experimental data from model predictions, no clean and robust procedure has been devised so far to account for this possibly significant problem.

### 3 Integral Data Testing

Here we make the distinction between *differential* and *integral* experimental data. A differential measurement attempts to measure the physical quantity of interest more or less directly. For instance, a fission time-projection chamber simply counts fission events to infer a fission probability or cross-section. As discussed above, such data are routinely used in nuclear data evaluations.

On the other hand, integral measurements do not provide a direct link to basic nuclear data, but may be simulated by combining different nuclear data of interest. The multiplication factor \( k_{\text{eff}} \) in a critical assembly measures the net number of neutrons produced from one
incoming neutron. If $k_{\text{eff}}$ is greater than one, the system is said to be critical, and sub-critical if less than one. This factor has been measured with great accuracy for some well-known critical assemblies such as Jezebel and Godiva (see Fig. 12). Transport simulations can be used to calculate $k_{\text{eff}}$ very precisely as well. The most relevant nuclear data that enter into its calculation are fission, elastic, inelastic and capture cross sections and prompt fission neutron multiplicity. Since multiple nuclear data sets can impact the result of this calculation, a measurement can only be used to constrain those data simultaneously and therefore provide cross-reaction correlations. However, it is difficult to use these experimental data points to constrain one cross-section uniquely.

Recent efforts have been made to include integral data as part of the evaluation process, in particular using the Total Monte Carlo method mentioned above [40]. Much work remains to be done though to understand how to properly use the results of this work, and to avoid biases in the final evaluations due to a particular choice of integral data sets.

### 3.1 Formatting and Processing

ENDF-6 is the standard format for storing, exchanging and processing evaluated nuclear data files. It is described at length in the ENDF-102 manual [44]. It is written in ASCII format, so is human-readable, and it is relatively easy to extract specific sections from a data file. However, it is rather complicated to fully master, due to the large number of flags and specific cases embodied in the format. Figure 13 depicts the beginning of the ENDF/B-VII.1 evaluated file for the neutron-induced reactions on Am-242. Important identifiers
repeated throughout the file include the MAT number that specifies the target nucleus, the 
MF number for the reaction type (cross sections, angular distributions, spectra, etc.) and
the MT number that corresponds to a particular reaction channel (total, capture, fission,
etc.). For instance, the section with MF=3 and MT=18 corresponds to the fission cross 
section data.

Figure 13: First ten lines of the ENDF/B-VII.1 evaluated nuclear data file for the neutron-induced reactions on Am-242. The MAT number (9546) specifies the target nucleus, and the following MT (1) and MF (151) numbers specify the reaction type and reaction channel. Here, MT=1, MF=151 is a special case corresponding to the
dictionary of the file, providing a description and index for all reactions included in the file.

A nuclear data evaluation file may contain the following information (among others):

- MF=1, MT=151: general information on the authors, comments on the evaluation process for different sections
- MF=2, MT=151: resonance region parameters
- MF=3: reaction cross sections (total, elastic, ...)
- MF=4: angular distributions for emitted particles for the different reaction channels
- MF=5: energy distributions (spectra) for emitted particles
- MF=33: covariance matrices for reaction cross sections
- MF=35: covariance matrices for energy distributions
- etc.

Of particular importance to our above discussion on uncertainties are the sections MF=30 through 40, which store the covariance matrices corresponding to various reaction channels, energy distributions, resonance parameters, angular distributions, etc. In the recent past, significant efforts have been made by the community to produce covariance matrices associated with the different reaction channels stored in the other MF sections of the file.

In this short lecture, we will not dwell any further into the intricacies of the ENDF format but rather refer the interested reader to the very well-written ENDF-6 manual [44].

The ENDF format was developed at a time when punch cards and old versions of the Fortran programming language were the norm. Processing codes such as NJOY [45], PREPRO [46],
AMPX [47] were all developed around this format. Those codes have evolved significantly over time, but still rely on the ENDF-6 format, relatively unchanged since its inception. This permanence is actually a very positive aspect of the ENDF-6 format, as it has been used over many years for many applications. However, it does have its drawbacks: lack of quality control, lack of documentation within data files, redundant information, difficulty to manipulate and plot, ... Modern technologies such as XML, HDF5, MySQL, Python, etc., could be used to extend the accessibility of the data, and improve the quality assurance of all data produced. The OECD/WPEC subgroup 38 [48] has been established to help design a framework for a new format. The Generalized Nuclear Data (GND) format, based on the extensible markup language XML, is a candidate to replace ENDF-6. Such a format could be used not only to store evaluated data as in the ENDF/B-VII library, but also to store experimental data as well as model calculations. Common analysis and plotting tools could then be developed for a large set of data types.

It is important to note that any new format and tools associated with it would have to be 100% compatible with the ENDF-6 format for some time to come so that the processing codes cited above can still be of use, given their ubiquity for application needs.

Evaluated nuclear data files stored in the ENDF format can rarely be used directly and have to be processed before they can be used in applications and transport simulations. The NJOY nuclear data processing system [45] provides this crucial link by reading ENDF files and producing point-wise and multi-group nuclear cross sections and related quantities that can be used in applications. It calculates Doppler-broadened cross sections, resonance self-shielding, heat production cross sections (KERMA factors), multi-group cross sections, and transforms ENDF data into various formats that can be used in transport codes such as MCNP.

### 3.2 Integral Data through Transport Simulations

Integral data are extremely useful to test the validity and accuracy of evaluated nuclear data. The ENDF/B-VII.0 and VII.1 libraries were and remain constantly benchmarked through the use of various integral data [49, 50, 51, 52, 53]. We provide here a few examples extracted from the main ENDF/B-VII.0 reference paper [49].

A first example concerns the use of bare and $^{238}\text{U}$-reflected unmoderated critical assemblies such as Jezebel and Godiva discussed above. The ratios calculation-over-experiment, commonly denoted as C/E, are shown for the $k_{\text{eff}}$ of several assemblies in Fig. 14. The experimental data and MCNP models come from the 2004 or 2005 editions of the Handbook of the International Criticality Safety Benchmark Evaluation Project (ICSBEP) [54]. The calculations were performed using the MCNP code (versions 4c3 or 5). Figure 14 shows a remarkable improvement from B-VI.8 (open squares) to B-VII.0 (closed squares) library.

Another interesting example of benchmarks is provided in Fig. 15, again taken from [49]. It represents the ratio of $^{238}\text{U}$ (n,2n) and $^{235}\text{U}$ (n,f) cross sections, plotted as a function of the ratio of $^{238}\text{U}$ (n,f) to $^{235}\text{U}$ (n,f) cross sections. Because $^{238}\text{U}$ (n,f) is a threshold fissioner, this ratio is somewhat representative of the hardness of the neutron spectrum. Flattop-25 and Topsy are fast critical assemblies with holes at different distances from the center,
3.3 Propagation of Uncertainties in Transport Simulations.

When comparing integral data to the results of transport simulations, it would be natural to check for the uncertainties in the measurements as well as theoretical predictions. While the experimental error bars are almost always provided, the opposite is true for simulated
data. Until recently the lack of evaluated covariance data was an obvious obstacle to providing theoretical error bars. This is not true anymore, and one should expect to see total (experimental+theoretical) error bars on most C/E plots.

The Total Monte Carlo (TMC) approach [40] provides a natural framework for propagating uncertainties from nuclear data. As mentioned above, it implicitly treats non-Gaussian distributions as well as non-linear effects. However, it is often limited to model parameter uncertainties, and includes only limited experimental information- albeit the model parameters have been previously tuned to available experimental data. Nevertheless, it remains a powerful tool to compute not only mean values but also probability distribution functions of integral data.

Another approach involves the use of covariance matrices associated with evaluated nuclear data. An efficient and powerful way to propagate uncertainties is to first find the eigenvalues and eigenvectors of a covariance matrix. In many cases, the number of significant eigenvalues is much smaller than the stored covariance matrix, due to the use of a model, with a limited number of parameters, to represent a large set of data. Also called Principal Component Analysis (PCA), this method can be used to speed up tremendously the propagation of uncertainties to integral data. When the number of eigenvalues is relatively small, the Polynomial Chaos Expansion (PCE) and Stochastic Collocation Method (SCM) are powerful transport methods to also cut down the computation time [55].

4 Concluding Remarks

To this day, nuclear data evaluations remain very relevant to the success of many nuclear applications and fundamental research topics. Traditional applications such as light-water reactors, for which a tremendous amount of feedback has been collected over the years, only little new data, if any, are needed. Even in this most extreme case however, updated nuclear data could be used to reduce safety and security margins, thereby improving the efficiency of those reactors. The advent and design of Generation III and IV reactors, small and modular reactors, as well as the continuing issue of nuclear waste management, have brought new requests and constraints on evaluated nuclear data and their uncertainties. Non-proliferation, safeguards, treaty verification, etc., remain important modern and active topics of research. Historically, approximate methods have been devised for solving different types of nuclear problems, always using approximate or average nuclear data. Moving beyond average nuclear data requires a revision of some of these methods and data.

In addition, the recent efforts by WPEC Subgroup 38 [48] to revamp the ENDF format align well with the modernization of the nuclear data evaluator toolkit. This is an important tool to modernize nuclear databases, improve the quality control of the evaluations, and facilitate their use. Interestingly, some applications may require nuclear data such as correlated data in fission [31] that simply cannot be stored in data files, but are better calculated on the fly in transport simulations. Even in this case, the improved and more detailed physics models that are used to produce correlated data can provide better insights into the evaluated values of many data.

The advent of new nuclear experimental facilities worldwide (FRIB, NFS, SPIRAL-2, etc., see for instance Ref. [56]) is sure to shake up our still limited knowledge on the fundamental
nucleon-nucleon forces, and provide some “stress tests” for existing theoretical models used in nuclear data evaluations. Common statistical nuclear reaction codes cited above all use many phenomenological models that have been tuned and fitted to existing data, mostly near the valley of stability. This is true for optical model and coupled-channel potentials, nuclear level densities, gamma-ray strength functions, fission barriers, etc. By exploring new regions of the nuclear landscape, the predictive power of those models will be tested, and new models or/and new parameterizations will be developed.

The significant increase in computing power has yet to be tapped in nuclear data evaluation work. While a fundamental treatment of the n-n interaction is prohibited for all but the lowest mass nuclei, significant advances in purely microscopic approaches such as Hartree-Fock-Bogolyubov have yet to be used routinely in nuclear data evaluation work. Paradoxically, the use of more fundamental physics models often tends to bring worse results for applications. This is in part due to the fact that evaluated data have already been adjusted, and that any tweak or modification to a particular cross section should be accompanied with similar changes to competing reaction channel cross sections simultaneously. This represents a tremendous obstacle that the nuclear data evaluation community is always hesitant to tackle, for many good (and bad) reasons.

Another venue for future improvements in evaluated libraries is the existence of correlations between isotopes and reactions. Nuclear fission data represent a particularly appealing example of such correlations. For instance, energy-dependent fission cross sections for a suite of isotopes are correlated through the multiple-chance fission mechanism in which neutron(s) can be emitted from the excited compound nucleus prior to fission, thereby linking the fission cross sections of different isotopes in the same isotopic chain (A, A-1, A-2, ...). Many other correlations exist in fission data, e.g., cross sections and angular distributions, which are not fully exploited in current evaluated libraries.

Ideally, one fundamental nuclear physics model would be capable of uniting all nuclear data observables. While this perspective will certainly remain elusive for a long time, any effort moving in this direction would improve evaluated nuclear data and the predictive power of nuclear models.

Finally, CIELO (Collaborative International Evaluated Library Organization) [57] is a recent project that aims at getting nuclear data experts together around common problems impacting significant materials and reactions, and hopefully at delivering an agreed-upon library of state-of-the-art nuclear knowledge. As a famous example, the value of the $k_{\text{eff}}$ for the Jezebel Pu critical assembly is simulated very accurately by all evaluated nuclear data libraries in the world. However, they all achieve a similar result for very different reasons. Fission, capture, elastic, inelastic cross sections, as well as fission neutron multiplicities and spectra, all differ from one library to the next. It can be concluded that compensating errors are ubiquitous and that evaluated nuclear data libraries have been somewhat adjusted to well-known and accurate integral benchmarks.

Some of the original motivations for developing CIELO are to identify reasons for discrepancies between libraries, to produce higher-quality evaluations, to minimize missing out on key measured differential/integral data, to increase peer-review, and to share work and responsibilities. This is definitely an important endeavor that expands on existing worldwide collaborations, and that can only further strengthen ties between an ever smaller number of
nuclear data evaluators.
References

[16] S. Hilaire, title here, these lecture notes.
[17] F. Gunsing, title here, these lecture notes.
[23] “The double-humped fission barrier,” S. Bjørnholm and J.E. Lynn, Rev. Mod. Phys. 52, No.4, 725 (1980).


