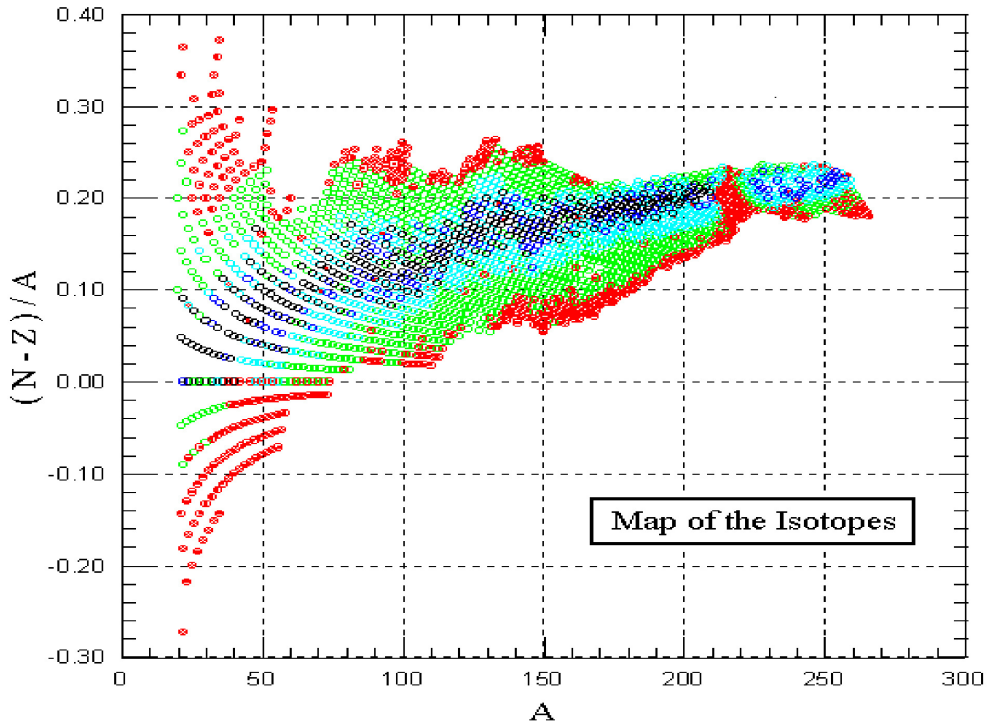
ANL/NDM-159

Covariance Matrices for Nuclear Cross Sections Derived from Nuclear Model Calculations

Nuclear Data and Measurments Series

Nuclear Engineering Division
Argonne National Laboratory



ARGONNE IS OPERATED BY THE UNIVERSITY OF CHICAGO FOR THE U.S. DEPARTMENT OF ENERGY OFFICE OF SCIENCE

About Argonne National Laboratory

Argonne is operated by The University of Chicago for the U.S. Department of Energy Office of Science, under contract W-31-109-Eng-38. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

Availability of This Report

This report is available, at no cost, at http://www.osti.gov/bridge. It is also available on paper to U.S. Department of Energy and its contractors, for a processing fee, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone (865) 576-8401
fax (865) 576-5728
reports@adonis.osti.gov

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor The University of Chicago, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, Argonne National Laboratory, or The University of Chicago.

ANL/NDM-159

Covariance Matrices for Nuclear Cross Sections Derived from Nuclear Model Calculations [a]

Donald L. Smith [b]
Nuclear Engineering Division
Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 60439
U.S.A.

November 2004

Keywords

METHODS. Nuclear cross sections, nuclear models, errors, covariances, Monte Carlo.

---

[a] This work was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under Contract W-31-109-Eng-38.

[b] *Author Contact*: 1710 Avenida del Mundo #1506, Coronado, California 92118-3073, U.S.A.; Tel - +1(619)435-6724; E-mail - Donald.L.Smith@anl.gov.

Nuclear Data and Measurement Series

The Nuclear Data and Measurement Series documents results of studies in the field of microscopic nuclear data. The primary objective is the dissemination of information in the comprehensive form required for nuclear technology applications. This Series is devoted to: a) measured microscopic nuclear parameters, b) experimental techniques and facilities employed in measurements, c) the analysis, correlation and interpretation of nuclear data, and d) the compilation and evaluation of nuclear data. Contributions to this Series are reviewed to assure technical competence and, unless otherwise stated, the contents can be formally referenced. This Series does not supplant formal journal publication, but it does provide the more extensive information required for technological applications (*e.g.*, tabulated numerical data) in a timely manner.

Table of Contents

- Information About Other Issues of the ANL/NDM Series ----- ----- p. 6
- Abstract ----- ----- ----- ----- ----- ----- ----- ----- p. 7
- Acknowledgements ----- ----- ----- ----- ----- ----- ----- p. 8

1. Introduction ----- ----- ----- ----- ----- ----- ----- ----- p. 9
2. Description of the Method ----- ----- ----- ----- ----- ----- p. 10
3. Mathematical Formalism ----- ----- ----- ----- ----- ----- p. 12
4. The Model Parameters ----- ----- ----- ----- ----- ----- p. 15
5. The Derived Quantities ----- ----- ----- ----- ----- ----- p. 17
6. Coding Issues ----- ----- ----- ----- ----- ----- ----- p. 18
7. Bookeeping and Shortcuts ----- ----- ----- ----- ----- ----- p. 19
8. A Suggested Test Case ----- ----- ----- ----- ----- ----- p. 20
9. Conclusions ----- ----- ----- ----- ----- ----- ----- ----- p. 21

- References ----- ----- ----- ----- ----- ----- ----- ----- p. 22
- *Appendix A* ----- ----- ----- ----- ----- ----- ----- ----- p. 23
- *Appendix B* ----- ----- ----- ----- ----- ----- ----- ----- p. 25

Information About Other Issues of the ANL/NDM Series

Information about all other reports in this Series is available from Argonne National Laboratory at the following Internet site:

http://www.td.anl.gov/reports/index.html

There one can also find a list of titles and authors, abstracts, and specific reports available for viewing or downloading as PDF files with no cost or restriction.

If any questions arise concerning the above-mentioned Internet site or the present report or any other report in this Series, please direct them to:

Dr. Filip G. Kondev
Nuclear Data Program
Nuclear Engineering Division
Building 208
Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 60439
U.S.A.

Tel: +1(630)252-4484
Fax: +1(630)252-5287
E-mail: kondev@anl.gov

Covariance Matrices for Nuclear Cross Sections Derived from Nuclear Model Calculations [a]

by

Donald L. Smith [b]
Nuclear Engineering Division, Argonne National Laboratory
Argonne, Illinois 60439, U.S.A

Abstract

The growing need for covariance information to accompany the evaluated cross section data libraries utilized in contemporary nuclear applications is spurring the development of new methods to provide this information. Many of the current general purpose libraries of evaluated nuclear data used in applications are derived either almost entirely from nuclear model calculations or from nuclear model calculations benchmarked by available experimental data. Consequently, a consistent method for generating covariance information under these circumstances is required. This report discusses a new approach to producing covariance matrices for cross sections calculated using nuclear models. The present method involves establishing uncertainty information for the underlying parameters of nuclear models used in the calculations and then propagating these uncertainties through to the derived cross sections and related nuclear quantities by means of a Monte Carlo technique rather than the more conventional matrix error propagation approach used in some alternative methods. The formalism to be used in such analyses is discussed in this report along with various issues and caveats that need to be considered in order to proceed with a practical implementation of the methodology.

---

[a] This work was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under Contract W-31-109-Eng-38.

[b] *Author Contact*: 1710 Avenida del Mundo #1506, Coronado, California 92118-3073, U.S.A.; Tel - +1(619)435-6724; E-mail - Donald.L.Smith@anl.gov.

Acknowledgements

The author is indebted to Dr. Arjan Koning, Nuclear Research and Consultancy Group, Petten, The Netherlands, for many fruitful discussions on this topic and for his ongoing efforts to demonstrate that the approach discussed in this report is indeed feasible and potentially practical. The author is also grateful to Dr. Filip Kondev, Argonne Nuclear Data Program, for posting the present document on the Argonne National Laboratory Internet site reserved for reports of the ANL/NDM Series.

1. Introduction

It is a reality of modern neutron cross section evaluation procedures that nuclear models play a pivotal role. While the quality of results derived from these models varies considerably depending on the reaction type, isotope, sophistication of the codes, availability of experimental data for comparison and benchmarking purposes, *etc.*, a distinct advantage of nuclear models is that the various partial cross sections automatically sum up to the total cross section, thereby leading to internal consistency of any evaluated file generated exclusively by using nuclear models. When the modeling is guided by reliable experimental data, where available, then the model parameters are constrained by these physical observations, and the overall quality of the file improves. In cases where explicit data are lacking, the experience and good judgment of the nuclear modeler, as well as systematic considerations, can serve as guides in constraining the choice of values for the parameters, often leading to reasonably reliable results, especially for the stronger reaction channels. While many of the contemporary evaluated files have been generated in large part by using nuclear models, the ability to generate covariance (uncertainty) information for these files is generally lacking, mainly due to absence of an agreed upon procedure for generating this information.

One approach that has been explored elsewhere is to provide estimates of uncertainties for the parameters used in the model calculations, and then propagate these uncertainties – with the aid of sensitivity parameters calculated from the nuclear model – to produce covariance information for the derived results. In fact, experimental data have even been fitted by such models and these fits used to generate evaluations [1]. Although this deterministic approach has been applied in a few cases, it is rather labor intensive. In this communication a different method is suggested, namely, a statistical one based on utilizing the prodigious power of modern computers and the well-known Monte Carlo technique.

2. Description of the Method

In simplest terms, the approach presented here proceeds as follows: The evaluator (nuclear modeler) first chooses a set of "best" parameters that will yield a complete evaluation for the particular isotope of interest. The selection of these "central values" will be guided by experimental data, experience of the evaluator, *etc.* It should represent the best effort that the evaluator is able to muster using the objective information at hand. The present approach is conceived, at this stage of its development, to apply only to the fast neutron region since those techniques used in evaluations of data for the thermal, resolved resonance, and unresolved resonance regions are all quite distinct. In principle, however, such a limitation is somewhat artificial, and consistent procedures for applying the present method to the thermal, resolved resonance, and unresolved resonance regions could be developed over time.

For convenience, a common energy grid should be selected for all considered reaction channels. Then, random variations are made simultaneously to all of the nuclear model parameters for each of a selected number of Monte Carlo "histories". For each history, a distinct collection of comparable derived values is calculated for all these reaction channels and selected energies. It is recommended that at least 1000 such histories (preferably more) be pursued in order to achieve a reasonable degree of statistical convergence (~3%). At the end of this sampling procedure, the collected Monte Carlo results are utilized in a statistical analysis which automatically generates a global covariance matrix that expresses the uncertainty in the calculated quantities, including cross correlations (or anti-correlations) between different energies and reaction channels. Using modern computer workstations (or even personal computers), a single history, corresponding to a complete evaluation with a given set of parameters, can generally be completed within a matter of minutes (or perhaps even seconds)! Thus, adequate statistical samples of such evaluations can be generated in a matter of several hours – or perhaps at most a day – of computer running time. Existing nuclear model code packages, with seamlessly linked modules, can be invoked as subroutines for a master "driver" code that prepares the input for each history and analyzes the accumulated results to produce the desired covariance matrix.

Critics will argue that the uncertainty associated with a particular evaluation generated using nuclear models cannot be limited solely to the effects of model parameter uncertainties. There is also the matter of model limitations and deficiencies as well as other considerations to account for. While this argument is certainly true conceptually, in fact it should be possible in a pragmatic way to account for these additional uncertainty sources in all the derived results reasonably well by their inclusion within the framework of parameter uncertainties, provided that these uncertainties are adequately posed. There appears, in the view of this author, to be no other viable alternative to generating in an objective way covariance matrices for evaluations derived directly from computer modeling. Certainly, if one accepts the idea that uncertainties in physical quantities calculated from model parameters ought to reflect the uncertainties in the underlying

parameters, the Monte Carlo approach suggested here – which in a sense is comparable to conventional error propagation – is inherently reasonable.

In this report, some of the details associated with implementing this approach are discussed. In practice, the utility of this method can be established only by demonstrating its applicability to several practical examples. However, no attempt has been made in the present work to carry out such an exercise because the requisite computational tools are not readily available to this author. However, as we shall see below, such an investigation is currently being pursued by another scientist who is experienced in the development and application of nuclear model codes (see Section 8).

One point is clearly evident to this author *a priori*: The covariance matrices generated by this approach will surely exhibit rather strong correlations. The reason for this is that relatively few parameters must carry the "burden" of generating a considerably larger number of derived results, *e.g.*, cross sections for a collection of reaction channels over a wide range of energies extending from perhaps a few tens of keV to several tens of MeV. This follows from the general observation that when a large number of values are computed using just a few parameters, the uncertainty correlations for the derived results are frequently large. A characteristic "stiffness" of evaluated files generated using nuclear models is unavoidable even if the correlations between the input model parameters are weak or non-existent (which, of course, is not entirely the case in reality). It is a fact of life that the nature of the covariance matrices generated by a selected evaluation process depends strongly on the input information used and the process employed in the evaluation itself. The present method, where nuclear models are involved, is no exception. Thus, the approach discussed here is in keeping with the broad notion that covariance matrices for evaluated quantities should be generated, where possible, by methods that are intimately consistent with the underlying evaluation procedures.

3. Mathematical Formalism

Let us suppose that p_ℓ is a typical parameter of the nuclear model used to generate an evaluated file. We shall then assume that there are L such parameters in total that exhibit uncertainty, i.e., ℓ = 1,L. We can represent these parameters by the vector **p**. Furthermore, assume that the symbol σ_i corresponds to a distinct physical quantity calculated using the nuclear model (embodied in a nuclear model code package) using parameters **p**. These can be a collection of cross sections for several reactions and energies, or can even be particle emission spectrum values, differential cross sections, gamma-ray spectra, *etc.* The list of these physical quantities that are to be included in the evaluation, as well as the energies, is established prior to the analysis described here. The list is finite, thus i = 1,n, and the collection can be represented symbolically by the vector $\boldsymbol{\sigma}$. A functional (one-to-one) deterministic relationship exits between **p** and $\boldsymbol{\sigma}$, *i.e.*, $\sigma_i = f_i(\mathbf{p})$ for i = 1,n. The relationship is complicated, however, and can be established only through the selected nuclear modeling procedure and extensive numerical calculations. For convenience, let $\mathbf{p}_0$ correspond to the evaluator's choice of "best" parameters for the model used to generate a particular evaluation, as indicated above. This choice is guided by experimental data, systematical considerations, experience, *etc.*, in the usual fashion. It is not envisioned that this should be a global parameter set such as one might find in RIPL [2], but rather a set of parameters "tuned" for optimal representation of the best available objective information for the specific isotope in question. Of course, in the absence of such an optimal "local" parameter set, global values might be used provided that adequate parameter uncertainties are assigned. The consequence of applying this choice of values for the parameter set is the collection of derived results $\boldsymbol{\sigma}_0$ that represents the final evaluation for the isotope in question. For convenience, we refer to these numbers as the "central" values for the purpose of analyzing uncertainties and generating a global covariance matrix associated with the evaluation.

Now, let us define "k" as the index that identifies a particular "history" in the ensuing Monte Carlo analysis. We supposed that k = 1,K. That is K histories are pursued in the statistical analysis. For each history, there corresponds a parameter set $\mathbf{p}_k$ and the derived set of evaluated results $\boldsymbol{\sigma}_k$. Each parameter set $\mathbf{p}_k$ is generated by varying all of the individual parameters relative to their central values in a manner to be discussed below. Each individual history, denoted by the index "k", corresponds to production of a completely new evaluation for the isotope in question, *i.e.*, it produces cross sections for all included reactions at all the selected energies, and, if desired, may also include differential and integral results as well as particle emission spectra. Of course, a subset of the complete evaluation could be examined without any loss of generality, if that is desired. The elements of the n x n global covariance matrix **V** of absolute uncertainties (encompassing all considered reactions and energies corresponding to a particular isotope) can be estimated in a straightforward manner from the following expression [3]:

$$V_{ij} = \langle(\sigma_{ik}-\sigma_{i0})(\sigma_{jk}-\sigma_{j0})\rangle \quad \text{for } i,j = 1,n\,. \tag{1}$$

The symbol <…> signifies averaging with respect to the results obtained from the collection of K histories. Thus, more explicitly,

$$V_{ij} = (1/K) \times \Sigma_{k=1,K} (\sigma_{ik}-\sigma_{i0})(\sigma_{jk}-\sigma_{j0}) \quad \text{for } i,j = 1,n . \tag{2}$$

Arguments over whether K or K-1 should appear in the denominator of Eq. (2) are of no significance since K should be as large as is practicable given the available computer resources, and certainly at least as large as K = 1000 (~3% statistical "accuracy" in the computed averages), as mentioned above. If a relative covariance matrix (*i.e.*, a matrix involving fractional errors) is desired, it can be obtained readily from the expression

$$R_{ij} = V_{ij}/(\sigma_{i0}\times\sigma_{j0}) \quad \text{for } i,j = 1,n . \tag{3}$$

If the correlation matrix is sought, it can be obtained directly from the expression

$$C_{ij} = V_{ij} /(V_{ii} \times V_{jj}) \quad \text{for } i,j = 1,n . \tag{4}$$

One might anticipate that

$$\sigma_{i0} \approx (1/K) \times \Sigma_{k=1,K} \sigma_{ik} \quad \text{for } i = 1,n . \tag{5}$$

That is to say that the mean values obtained from analysis of Monte Carlo sampling results should lie rather close to the central values as defined above. If the parameter uncertainties are relatively small, and consequently the spreads in derived values obtained by Monte Carlo simulation are also relatively small, then the assumption expressed in Eq. (5) is probably reasonably valid. However, Eq. (5) need not be unconditionally true due to the largely non-linear relationships between the derived results and the model parameters. In any event, mean values and central values should be compared as part of the analytical process to insure that nothing has gone seriously wrong in the procedure.

Equations (1) - (5) offer a rational approach – in a statistical sense – to the quest for the appropriate covariance matrices to apply for evaluations based entirely on nuclear modeling. Furthermore, by resorting to Monte Carlo simulation one avoids the difficulties associated with determining a very large number of sensitivity parameters numerically (actually they are approximations to partial derivatives), although explicit sensitivity information can be extracted from this analysis if desired. All the needed variances (or standard deviations) are produced automatically along with correlations between distinct energies, various reaction channels, *etc.* In short, by implementing this approach we work the computer very hard and our brains less so. The results obtained using the present method can be cast into the appropriate selected evaluated data formats, and our problem of producing viable covariance matrices for evaluations generated from nuclear models is solved, at least for the fast-neutron region. The global matrix (or the various sub-matrices) obtained in this manner for the fast neutron region can eventually be augmented, as required, to include uncertainties for the thermal, resonance, and unresolved resonance regions. Each distinct region can be addressed according to the

dictates of the applicable evaluation methodology. Cross-energy-region covariances can then be introduced, most probably in an ad hoc manner initially, in order to yield complete covariance matrices for the entire general purpose “vertical” evaluation.

But, it is widely acknowledged that “God” (or the “Devil” for that matter) can be found in the details. Improperly handled, such numerical details could doom a conceptually reasonable approach to certain failure. So, in the next sections some of these details are discussed and suggestions as to how the various pertinent issues might be addressed are offered.

4. The Model Parameters

It was noted above that a collection of parameters **p** is used to characterize the model (or computational process if you will) that generates an evaluation denoted symbolically by the vector $\boldsymbol{\sigma}$. Among these parameters are particle potential radii, well depths and diffuseness, discrete level parameters, level densities, *etc.* These parameters are of two types: continuous and discrete.

Uncertainties for continuous parameters can be represented by probability density functions or, more commonly, by mean values and standard deviations. If the uncertainties are modest (*e.g.*, less than 20%), then these distributions can be adequately approximated by Gaussians [3]. If the uncertainties are larger, and the physical quantities are defined to be non-negative, then the lognormal probability function should probably be used [4]. In any event, the sampling of parameter space can, in general, be carried out in accordance with the usual methods used to sample random variables from distributions. However, it is well known that the parameters that constitute the numerical embodiment of a nuclear model cannot be treated as completely independent random variables. There are physical constraints between these parameters to consider, *e.g.*, such as the one introduced by the well known "$Vr^2 \approx$ invariant potential well strength" relationship for nuclear potentials [5]. These constraints can lead to non-negligible correlations. Thus, sampling to generate random parameter vectors $\mathbf{p}_k$ for the Monte Carlo exercise must be carried out with great care to take such correlations into account. When correlations do exist, the individual parameters of this vector may not be varied completely independently. One way to carry out the sampling, at least when the uncertainties are modest, might be to perform a linear transformation from real parameter space to a pseudo space of comparable parameters where the correlations essentially vanish. Values for these pseudo parameters could be selected independently at random (without concern for correlations) and then transformed back again to realistic parameters for the purpose of calculating physical quantities in the Monte Carlo trial. The details of how to do this properly will probably entail a fair amount of additional investigation and such an effort should be undertaken eventually.

The influence of parameter correlations, while ultimately of considerable importance, is nevertheless a detail that can be examined later in the development of the present method. For simplicity, the concept discussed here should be demonstrated at the outset by assuming that all the continuous variable parameters of the model are uncorrelated. In fact, it would be interesting to eventually compare the results obtained with and without the inclusion of correlations in the model parameter set. Of course, even if the model parameters are treated as uncorrelated, correlations will emerge for the derived cross sections, *etc.* In passing, it should be noted that such a Monte Carlo approach has been used with considerable success by the nuclear astrophysics community, especially by the ORNL group [6]. The "parameters" of their analyses are stellar reaction rates. Uncorrelated uncertainties are assumed for these values. The "model code", in this case, is a model of stellar evolution. The derived results are mean values for stellar isotopic abundances and their associated uncertainties.

Another category of model parameters is the discrete category, *e.g.*, spins, parities, and those discrete parameters specifically associated with the pre-compound model. How should uncertainty be handled here? Consider an example. An important low-lying level might have either spin/parity $1/2^-$ or $3/2^-$. If we don't know which it is we could assign 50% probability to each possibility, and for every individual history select one or the other based on a binary "coin toss" approach; this is not hard to do with random number generator routines. Finally, one should take care to insure that the uncertainties assumed for all the various parameters of the model correspond to completely equivalent degrees of confidence for each of them [3]. Simply stated, it is inappropriate to assume uncertainty corresponding to a one-sigma error bar for one parameter while assuming a three-sigma error bar uncertainty for another one.

5. The Derived Quantities

It has already been mentioned that σ can symbolize a wide variety of derived physical quantities generated by nuclear model calculations. So, in practice there exists a bookkeeping problem to deal with. In all likelihood, the first covariance files that would be generated by this method would correspond to integrated, energy-dependent cross sections for a selected set of reaction channels addressed by the model, and for a collection of chosen energies. The covariance matrix for a particular reaction, but corresponding to all the selected energies over the range considered, would correspond to a sub-matrix of the global covariance matrix described above. Correlations between elements of this sub-matrix will appear, and correlations or anti-correlations (negative correlations) between this sub-matrix and other sub-matrices for other reactions would also be generated. The evaluator's dream of being able to ascertain correlations between various reaction channels would be realized because evaluations generated by nuclear models produce, simultaneously and in a unified fashion, values for cross sections and other observables corresponding to all considered reaction channels and energies. The power of this approach is impressive, but the evaluator should not lose sight of the underlying simplifying assumptions of the model, nor of the fact that the uncertainty estimates and correlations assigned to the model parameters may, by necessity, be somewhat *ad hoc*. As long as this point is kept clearly in mind by the evaluator, and the inherent deficiencies and limitations of the process are made known to the user community to the extent possible, it would appear that the advantages of this approach outweigh the disadvantages. Error bands corresponding to the derived covariance matrices should be plotted along with the central values and experimental data, where available, to see if the obtained standard deviations are reasonable. In particular, in cases where data exist, even if sparse, the parameter errors should be chosen so that these error bands are reasonably consistent with the observed scatter in the available experimental data.

6. Coding Issues

To implement the approach discussed above requires the development of a driver/controller master computer code that utilizes a chosen suite of nuclear modeling codes in a subordinate mode, *i.e.*, as subroutines. The driver code needs to sample model parameter space and carry out all the bookkeeping required to perform the statistical analysis describe here, to produce the covariance matrix, and to cast the results into appropriate formats. This driver program calls upon as its "slave", the chosen suite of nuclear model codes a total of K times, thereby generating a collection of results for the ensuing statistical analysis as described above. An additional code module might be required to assemble the covariances from the driver code and produce evaluated library files in the chosen evaluated data format, *e.g.*, the ENDF format that is widely used in applied nuclear science [7].

7. Bookeeping and Shortcuts

The output from contemporary nuclear model codes is generally quite extensive. It taxes the mind to conceive of dealing with 1000+ such sets of output in the forms currently encountered. Consequently, for the approach discussed here to be practical, it would be appear necessary to avoid collecting all this information explicitly. Rather, the sums associated with the covariance matrix elements, *i.e.*, those expressed in Eqs. (1), (2), and (5) above, should be generated as running sums, on the fly, with no intermediate values retained for the derived parameters. What is the point of keeping such information anyway? The only possible justification for retaining intermediate information would be for use plotting and possible some other non-essential statistical purposes. In fact, this approach could be used to test for statistical convergence of the computed values for those average quantities <…> which are defined above, and thus offer some indication as to how many histories "K" are required to achieve a reasonable degree of convergence for the matrix elements. Francis Perey once mentioned to this author that there is no point in discussing the uncertainty in an uncertainty [8]. Nevertheless, it is commonly accepted that a certain degree of stability (robustness) is desired for the values one assigns to elements of a covariance matrix. Numerical criteria could be established to guide the Monte Carlo procedure toward a desired level of stability for the computed values <…>. One criterion might be a comparison between the average values for cross sections and other selected observables and the selected central values (*e.g.*, see Appendix A). It should be noted here that for every history one obtains values for each of the various computed physical quantities, even those for which the cross sections are very small. Thus, the degree of "smallness" or "largeness" of the computed values for cross sections or other observable quantities derived from the model is not an issue that influences the statistical outcome.

The last point to be made is that one needs to consider which types of probability distributions are to be used in sampling the nuclear model parameters within their respective ranges. The statistical considerations involved in selecting these distributions are discussed in detail in the monograph by Smith [3] and again to a limited extent in Appendix B below.

8. A Suggested Test Case

A reasonable test case to investigate at the outset might involve performing an evaluation for a structural material isotope such as ^{56}Fe. There exist considerable experimental data for this material, corresponding to several reaction channels. The major cross sections to consider below 20 MeV are: neutron total, neutron elastic and inelastic scattering, neutron capture, (n,2n), (n,3n), (n,p), (n,np+pn), (n,α), (n,nα+αn), and possibly (n,^{3}He). One could select a reasonable number of energies spanning the region from about a few tens of keV to 20 MeV to form a common grid, and then apply the present methodology to determine the covariance matrices. Whether or not one believes the results for this early test case in great detail, it nevertheless would be interesting to show how correlations and anti-correlations between the evaluated results emerge for a particular reaction at various energies or between the distinct reaction channels at specific energies. The author has learned that Arjan Koning has utilized his model code suite TALYS recently to investigate the approach described in this report, and he has considered ^{56}Fe as a test case for his analysis [9]. The results of Koning's work are preliminary and they will be published later when his investigation is complete. Thus, they are not discussed in this report. At this time, however, it can be mentioned that Koning finds the present approach to be entirely feasible. Furthermore, he has demonstrated that this method offers considerable promise as a practical tool for use in generating covariance matrix information for evaluated nuclear data libraries.

9. Conclusions

A conceptually simple method for generating covariance matrices for evaluations based on nuclear modeling calculations is presented. The essential idea is that all the nuclear model parameters are varied randomly in a Monte Carlo exercise involving the generation of multiple evaluations. The covariance matrix elements are derived from the sampling results using equations based on the statistical definition of a covariance matrix. The actual application of this method is more complicated. It requires having a suite of nuclear model codes that can be called upon as subroutines by a driver code that randomly varies the model parameters and also performs the statistical bookkeeping tasks required to evaluate the covariance matrix elements. Early tests of the present approach by Arjan Koning (NRG-Petten) have demonstrated that this approach is both feasible and reasonably practical [9]. However, extensive further work is required to determine reasonable uncertainties and error correlations for the nuclear model parameters that are to be used for the generation of specific general purpose isotopic evaluated nuclear data files.

References

[1] Toshihiko Kawano *et al.*, International Conference on Nuclear Data for Science and Technology, Santa Fe, New Mexico, 26 September – 1 October 2004 (to be published as an AIP Conference Proceedings document).

[2] Reference Input Parameter Library (RIPL). These parameters can be obtained from the National Nuclear Data Center (NNDC), Brookhaven National Laboratory, NY, U.S.A, at the following Internet site: http://www.nndc.bnl.gov.

[3] Donald L. Smith, *Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology*, American Nuclear Society Press, LaGrange Park, Illinois, U.S.A. (1991).

[4] Donald L. Smith, Dmitri G. Naberejnev, and Laura A. Van Wormer, "Large Errors and Severe Conditions", *Nuclear Instruments and Methods in Physics Research* **A488**, 342-361 (2002).

[5] Donald L. Smith and Peter T. Guenther, "Covariances for Neutron Cross sections Calculated Using a Regional Model Based on Local-Model Fits", Report ANL/NDM-81, Argonne National Laboratory, Argonne, IL, U.S.A. (1983). Document available as a PDF file at the following Internet site: http://www.td.anl.gov/reports/ANLNDM81.html.

[6] Michael S. Smith, "Monte Carlo Approach to Determine the Impact of Nuclear Reaction Rate Uncertainties on Nova Models", Eighth International Symposium on Nuclei in the Cosmos, Vancouver, BC., Canada, 19-23 July 2004.

[7] Evaluated Nuclear Data File (ENDF) format, National Nuclear Data Center (NNDC), Brookhaven National Laboratory, NY, U.S.A. Information available at the following Internet site: http://www.nndc.bnl.gov.

[8] Francis Perey, Oak Ridge National Laboratory, Oak Ridge, TN, U.S.A. (Private Communication, *ca.* 1981).

[9] Arjan Koning, Nuclear Research and Consultancy Group, Petten, The Netherlands (Private Communication, 2004).

[10] Soo-Youl Oh, Nuclear Data Evaluation Laboratory, Korea Atomic Energy Research Institute (KAERI), Yusung, Taejon, South Korea (Private Communication, 2004).

Appendix A

In the main body of this report, an approach for generating covariance matrices in situations where nuclear models are used to produce evaluations is suggested. In that discussion, it was assumed that central values for the evaluated quantities were generated *a priori* (before the sampling procedure is initiated) by the modeler/evaluator, based on visual comparisons to experimental data (where available) supplemented by experience and the best judgment of the practitioner. This is the contemporary approach followed by most evaluators. These central values were denoted in the main body of this report by the subscript "0", *e.g.*, σ_{0i} for a specific quantity being evaluated. Of course, there is the additional requirement of providing estimates for the ranges of all parameters being sampled and, if possible, even probability distributions to govern the sampling procedure (see Appendix B).

However, there are many instances where there are no experimental data or other objective information available *a priori* to guide the evaluator in providing central values. The best that can be done by the evaluator is to resort to chosen ranges of parameters based on various considerations such as systematics, experience with neighboring nuclei, *etc.* Under these conditions, Soo-Youl Oh has pointed out that an alternative formulation of the present method is required, as discussed below [10].

The approach described here is similar in many ways to the one discussed in the main body of this report. In fact, the same nomenclature and equations can be used, *i.e.*, Eqs. (1) – (5) in Section 3. However, the differences are both conceptual and tactical. First, one determines the central values σ_{0i} by applying Eq. (5) from Section 3. No approximate sign ("≈") is used. Consequently, in principle these central values must be calculated first from averages of sampling results before the covariance matrix can be generated. Then, and only then, can one apply Eqs. (1) – (4) from Section 3. In short, the central values are estimated *a posteriori* based on averaging the results obtained from repeated sampling of all the parameters, as described earlier. Once these average values are determined, then the elements of the covariance matrix can be evaluated. This would seem like a lot of extra work but, in fact, it is not really that difficult because the advantage of the following trick can be exploited.

First, it is required to write both Eqs. (1) and (2) from Section 3 in a somewhat different form:

$$V_{ij} = \langle\sigma_{ik} \times \sigma_{jk}\rangle - \langle\sigma_{ik}\rangle \times \langle\sigma_{jk}\rangle \quad \text{for } i,j = 1,n\,. \tag{A.1}$$

The symbol <…> signifies averaging with respect to the results obtained from the collection of K sampling histories. Thus, more explicitly,

$$V_{ij} = (1/K) \times \Sigma_{k=1,K}\,(\sigma_{ik} \times \sigma_{jk}) - (1/K^2)\,[\Sigma_{k=1,K}\,\sigma_{ik}\,] \times [\Sigma_{k=1,K}\,(\sigma_{jk}\,] \quad \text{for } i,j = 1,n\,. \tag{A.2}$$

It is clear from these two equations that the sequence of samplings actually needs to be performed only once since running sums of the quantities σ_{ik} , σ_{jk} , and $\sigma_{ik} \times \sigma_{jk}$ can be evaluated as the sampling exercise progresses.

In short, the approach described here can be applied in cases where there is insufficient information available for the evaluator to specify central values *a priori*. The central values, along with their covariance matrices, are derived from a statistical sampling treatment based only on presumed knowledge of the parameter ranges and possibly of their probability distributions.

Appendix B

Two additional aspects of this subject are worthy of a brief discussion. They fall within the two distinct categories discussed briefly below:

B.1 Probability Distributions for Uncertain Parameters

According to the Principle of Maximum Entropy upon which Bayesian statistics draws widely, there are three reasonable possibilities for the model parameter probability distributions to consider in the context of the present situation [3].

1. If one has no further idea about the nature of a particular model parameter "p" other than its range, *i.e.*, that it falls within the parameter limits (p_{low},p_{high}), then one should assume a uniform (constant) probability distribution over this range and zero outside. That is, each values of the parameter within the selected range is to be treated as equally likely while no values outside the range need be considered.

2. If one has an estimate of the standard deviation Δp (error or uncertainty) in a parameter, as well as a decent estimate of the central value, p_0, and if the uncertainty $\Delta p << p_0$, then one should assume a normal (Gaussian) distribution for the parameter p with Δp as the standard deviation and p_0 as the mean value.

3. If the same conditions prevail as in #2 except that p is known to be inherently positive and Δp is NOT small, then one should assume a lognormal probability distribution for p with mean value p_0 and standard deviation Δp [4]. This prevents negative values for p from being generated during sampling exercises. Furthermore, it is well known that the lognormal probability distribution approaches the normal distribution in the limit of small uncertainties [3,4].

B.2 Correlations in Parameter Uncertainties

Considerable concern over the issue of correlations between parameters has been expressed in numerous private communications to this author. Indeed, this may be a problem, but it is suggest that it should first be established just how serious a problem it really is before exerting a lot of effort to refine our understanding of the nature of these correlations. First, even if the parameters are sampled in an entirely uncorrelated fashion there will still be strong correlations observed in the various derived values (cross sections, angular distributions, particle emission spectra, *etc.*). These correlations are introduced courtesy of the nuclear model itself. The issue that should concern us is that of establishing what effect the assumption of correlations – any *ad hoc* correlations – might have on the computed results. It would seem reasonable to start out with modest goals in mind. For example, one might compare results obtained with no correlation between the potential radius (r) and potential well depth (V) with those obtained by assuming very strong correlations between these two parameters, *e.g.*, such as would maintain Vr^2 nearly an invariant. Such *ad hoc* tests could serve to provide a better understanding of the

impact of correlations on the process of generating mean values and covariance matrices by Monte Carlo sampling. A healthy dose of common sense is also in order in this regard. For example, the parameters of the potential wells for α-particles, as required to calculate (n,α) cross sections, will probably not be correlated very strongly – if at all – with the parameters for the proton potential wells, as required to calculate (n,p) cross sections. However, the parameters chosen for the incident neutron potential well will affect all the cross sections and thus introduce significant correlations between the (n,α) and (n,p) results.