MAJOR QUESTIONS ABOUT DERIVATION OF VARIANCE-COVARIANCE INFORMATION
FOR NUCLEAR DATA EVALUATIONS*

R. W. Peelle

Oak Ridge National Laboratory
Oak Ridge, Tennessee 37830
United States of America

MAS

The uncertainties in and correlations among some evaluated nuclear data are now evaluated to permit estimation of data-related uncertainties in the outputs of neutronic calculations and to focus data improvement efforts. Questions are discussed that arise in trying to obtain adequate numerical files of variance-covariance uncertainty information. These involve 1) discrepant data, 2) experimental data with incompletely reported uncertainties, 3) uncertainties in nuclear model results, 4) uncertainty data for the resonance regions and for angle and energy distributions, and 5) the role of integral data in nuclear data evaluation. The question also arises whether files of uncertainty data designed for technological applications can suffice to represent past knowledge in an evaluation that includes new data. Directions are indicated toward resolving these questions.

*Research sponsored by the Division of Research and Technology
U. S. Department of Energy, under contract W-7405-eng-26 with the
Union Carbide Corporation.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. It does not represent the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed 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 imply endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

"By acceptance of this article, the publisher or recipient acknowledges the U. S. Government's right to retain a non-exclusive, royalty-free license in and to any copyright covering the article."

E. & B.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Introduction

In developing a file of nuclear data for engineering applications, initial priority is given to sufficient completeness and to obtaining results from neutronics calculations that are comparable to those from experiments. At a later stage, the need is felt to quantify the quality of the codified data to permit the data-related uncertainties in computed results to be estimated and to help identify candidates for measurement programs. Without quantified statements of uncertainty and correlation there is no guide to the quality of a data file other than comparisons with integral experiments that are also used to judge methods and are sometimes used to adjust cross section data.

Efforts toward adequately complete characterization of uncertainty in evaluated nuclear data have been seriously undertaken at a modest level for a decade. These efforts are bearing fruit in development and use of improved evaluation techniques,^{1,16} inclusion of extensive uncertainty files in practical evaluated data files,² and in the publication of many uncertainty analyses for neutronics calculations.³

Some definitions are required to clarify what is meant by evaluated nuclear data uncertainty information. Cross section data elements of interest here, $s_1, s_2, \dots$, have exact true values that remain unknown to us. A multivariate density function $f(s'_1, s'_2, \dots)$ characterizes our knowledge of these values based on the research that has been performed, such that we believe the probability that the true value of the evaluated data vector lies in the element $ds' = ds'_1 ds'_2 \dots$ at $\underline{s}'$ is just $f(\underline{s}') ds'$. Then the evaluated cross section

$$s_i \equiv \int \dots \int s'_i f(\underline{s}') \underline{ds}'$$

is the expected value (mean) of this density function, the variance or mean square "error" or uncertainty is

$$\text{var}(s_i) = \sigma_i^2 = \int \dots \int (s'_i - s_i)^2 f(\underline{s}') \underline{ds}' ,$$

and the covariance between two data elements is

$$\text{cov}(s_i, s_j) = \int \dots \int (s'_i - s_i)(s'_j - s_j) f(\underline{s}') \underline{ds}' .$$

The effort up to now has been to parameterize the complex multivariate density function by its moments of first and second degree; we need not assume the density function is normal in form. This focus on the second moment of the density function is convenient because of the easy rules for combining variances when a quantity can be considered a linear function of the variables of the density function.

Note that if the evaluated cross sections must be represented directly as a continuous function of energy, some approximation must be employed to prevent codified uncertainty data from being intractably voluminous; in any case we know very little about the value of the cross section at a "point" in energy. For energy dependent "smooth" cross sections, piecewise constant covariance quantities are therefore employed in the ENDF/B system.

If the evaluated data are in multigroup form, then uncertainty (or "covariance") data can be written as the matrix of variance and covariance elements:

$$v_{ij} \begin{cases} \text{var}(s_i), & i = j \\ = \\ \text{cov}(s_i, s_j), & i \neq j \end{cases}$$

Based on first order theory, the variance of a computed response R induced by the nuclear data $\underline{s}$ from which R is computed is given by

$$\text{var}(R) \approx \sum_{ij} \frac{dR}{ds_i} \frac{dR}{ds_j} v_{ij}.$$

If the density function goes to zero not too far from the evaluation $\underline{s}$, the central limit theorem assures that computed results depending on many cross sections will have nearly normal density functions.

Certain issues are sidestepped by development of nuclear data uncertainty files along the idealized lines implied above. Systematic evaluation failures or biases are not treated, yet the formats chosen to represent the evaluated cross sections themselves mandate both bias and methods-related uncertainty. The non-uniqueness of evaluated cross sections occurs because the various evaluation efforts have not utilized exactly the same knowledge bases. Finally, it is apparent that judicious approximations are required to identify, obtain, and codify the significant portion of the defined variance-covariance information for an entire cross-section set.

The sections below discuss some of the major questions that now arise in developing and codifying nuclear data uncertainty information; the initial intractable problem of estimating uncertainty data retrospectively for existing evaluated data is set aside here because evaluated data and their uncertainties are intrinsically the results of a single process.

Incompletely Reported Experimental Data

Our nuclear data evaluators seek to combine the world's knowledge of the relevant cross sections, yet the experimental results with which evaluators must work are

infrequently reported with enough attention to their needs. Information must be included to enable update of the results themselves to compensate for revised standards etc. More relevant here, the published variance-covariance information must be sufficiently explicit that the corresponding matrix intended by the author for the presented results can be generated without ambiguity. When only a few results are given, the matrix itself can be reproduced,⁵ while in more complex cases the publication of sufficiently complete information has not posed insuperable difficulties.⁶ Note that it is insufficient in general to just quote "systematic uncertainties" for the various experimental points -- the covariance elements are rarely well defined that way.

Almost as useless to an evaluator as ambiguously specified variance and covariance elements is the unjustified assertion by the experimenter of very small systematic uncertainties. Taken at face value, the claimed uncertainties might require the evaluator to neglect all other experiments though they appear to have been conducted with comparable care. Evaluators need to be bold in revising claimed uncertainties under such circumstances, even though severe criticism is inevitable. The next section covers the case where experiments appear to be discrepant because the assigned uncertainties are too small.

Conflicting Data

We intuitively abhor ambiguity, so tend to be concerned about differences even when the range of experimental results is narrower than the claimed uncertainties. Data points appear to differ when their "error bars" no longer overlap, yet in the comparison of any two large consistent data sets one should observe very many points that differ by twice the combined uncertainties. Real discrepancy is identified by differences three or more times as large as the combined uncertainties; where more than two sets are being combined, chi-square tests are helpful indicators of discrepancy though one must remember that the tabulated probabilities are not quantitative unless the density functions of the values being considered are normal in form. Since the scatter among experimental results can scarcely be ignored in the assignment of uncertainty, methods like SUR⁷ or SURP⁸ based entirely on data scatter may sometimes be justified.

The relatively rigorous methods of data combination based on various forms of least-square theory⁷ can aid combination of basically consistent data sets and perform the propagation of uncertainties. However, such methods of data combination are even more helpless than

our intuition in the face of real data discrepancies. Figure 1 lists methods used by evaluators to deal with discrepant inputs. Normative symbols indicate this author's judgement about their typical suitability. Those with the plus (+) symbol are recommended, those in parentheses are discouraged, and the remainder are suggested to be used with caution. The recommended procedures are time consuming but are much preferable to arbitrary selection or weighting of data, or to propagation of input uncertainties without concern for consistency. In this figure "poor documentation" includes failure to give complete variance/covariance information as discussed in the previous section.

06

The schematic resonance parameter fit at the bottom of Fig. 1 was chosen to illustrate that the misfit near the center of the figure should primarily affect the uncertainties of those parameters that represent the center peak. The same principle applies when the basis chosen for the parameterization is the cross section set itself. Where discrepancy is general between data sets rather than concentrated as in Fig. 1, one may without fear of serious error expand the uncertainties from least-squares procedures to "cover" the discrepancies. (However, a rigorous justification of this common technique cannot be given.)

In the presence of real data discrepancy, the role of the evaluator is here seen as the strong one of updating input quantities given by the experimenter. This uncomfortable role requires consultation with the providers of data and documentation of actions taken and the reasons for them.

Uncertainty in the Results From Nuclear Models

Particularly in the energy range above 14 MeV, evaluators often use nuclear reaction models with some parameter adjustment or verification by comparison with any available direct experimental results. The parameters of these models have generally been obtained from fitting of experimental results for theoretically related quantities, but for a variety of reasons there is little quantified information available on the uncertainties in and correlations among the model parameters. Moreover, even for cases where the general class of nuclear models to be used is well agreed upon, the models are approximate representations of the physics and results are expected to deviate from true cross sections even when there is enough direct data to adjust parameters for a particular case.

A few authors have addressed the uncertainty of nuclear model outputs and have achieved some success.⁹ To this author, it appears that a general treatment of

6

uncertainty in model calculations would require 1) obtaining a variance-covariance matrix of the model parameters from analysis of the experiments that led to the parameter selection, including whatever specific data was applied to adjust the model for the case of concern, 2) propagation of the parameter variance-covariance matrix to give the uncertainties and correlations for all the cross sections computed with that model. Since bias caused by systematic inadequacy of a model cannot be handled by present uncertainty treatments, the first step may include supplementing the "pure" physics model with additional parameters needed to represent known systematic model failures. Experience concerning the applicability of a global model to a specific nuclide must be considered so that available data for that particular nuclide will carry appropriately high weight. More examples are needed of techniques that allow one to estimate and codify the patterns of uncertainty in model outputs.

07

The Resonance Region

In the resonance cross section region the evaluator is challenged with detailed and complex cross section behavior not always clearly seen in experiments, but is aided by some knowledge of the formulae that should be applicable to represent this behavior.

For cases in which only a few resonances are important, the existing data can in principle be fit and the variance-covariance matrix of all the resonance parameters can be entered directly into an evaluated file to allow whatever uncertainty propagation is required.¹⁰ This direct approach is rarely exercised because the resonance parameter uncertainties usually obtained from fitting codes are unrealistically small. This effect is believed to occur because resonance analysis codes up to now have typically not permitted the user to include the systematic covariances among the input data that should dominate the parameter uncertainty when many data points combine to determine the value of the parameter itself. (The theorem on which rests the calculation of the variance-covariance matrix of least-squares fitting parameters requires that the complete variance-covariance matrix of the input data be included.) This simplest case is solved in principle; it remains to obtain experience with more advanced resonance analysis codes that accept input covariance data.¹¹

For nuclides that have many resonances that are of likely importance, such as ⁵⁶Fe or ²³⁸U, an almost

prohibitively large variance-covariance matrix for resonance parameters would have to be developed if the above solution were to be applied, and obtaining the needed sensitivity coefficients might be uneconomic. In any case, since the long-range uncertainties important to any broad-group cross section set must depend on the correlated uncertainties in the underlying experimental data, it appears that the covariances among values of corresponding parameters of the various resonances could be evaluated and codified by energy region just as are smooth cross sections. This hypothesis has not been verified, so what is needed is derivation of analytical results for simple cases, verification using numerical examples, and finally application of the resulting technique. Through such a development we could hope to obtain the uncertainties in shielded cross sections throughout the resolved resonance regions.

08

In evaluations where a portion of the resonance region is represented with "statistical" resonance parameters, additional problems impede progress toward achieving the ability to obtain appropriately complete uncertainty representations.¹² So little experience is available and there is so much question about the evaluation techniques used that it must suffice here to ask whether the local cross section fluctuations expected within any unresolved resonance treatment should be represented in the uncertainty file, and to note that the techniques to be developed must take into account the evaluation method used to obtain the statistical parameters and the uncertainties in those methods as well as in the data employed.

Angle and Energy Distributions

There is little experience known to the author in parameterizing the uncertainties in angle, energy, or joint angle-energy distributions. Since these distributions are a function of neutron energy, the large quantity of data that appear to be required has inhibited trials and format definitions.

In the case of secondary angle distributions evaluated in terms of their Legendre coefficients, the natural solution will be to tabulate piecewise constant uncertainties in and among these coefficients as a function of incident neutron energy for a few low-order coefficients. If the evaluated file instead contains tabulated angular distributions, similarly defined quantities based on symmetric and asymmetric components of the anisotropy may be appropriate parameters. The main research required before implementation of such representations is a demonstration that covariance information in these forms can readily be processed to obtain the quantities needed for uncertainty

propagation in transport calculations.

The problem in representing the uncertainties in secondary energy distributions such as fission spectra is that not only are the parameters uncertain but also the forms of the spectra are not very well known. Therefore, the variance and covariance elements that could readily be obtained as a function of incident neutron energy would likely underestimate the uncertainties at very high secondary neutron energies, for instance. Nevertheless, codifying the uncertainties in the parameters would be a valuable step. Marable et al.¹³ have computed sensitivity coefficients of some integral experiment results to the mean and the variance of the secondary energy distribution; the uncertainties in these quantities could be sought and tabulated. Further progress in this area awaits additional practical demonstrations.

The Role of Integral Data

Numerous examples have been given¹⁴ of inclusion of integral with differential data to allow refinement of cross sections. One can at least improve the ability to calculate the results of integral experiments similar to those used in the data "adjustment." Such refinement is sound in principle provided that full covariance data for the integral experiments and calculations are included, the cross section set being adjusted is sufficiently complete, and calculational methods errors ("biases") can be assumed to be negligible or to cancel out between the calculations of the integral experiments and practical systems. In any such technique, care must be exercised that the integral data used to adjust the cross section set not have already been used to guide the evaluation of the input differential data.

09 Examples for which cross section adjustments might readily be attempted using a basis of cross-section variables too incomplete to allow broad application of the results include (1) multigroup sets for shielding applications with broad energy groups and low-order angular expansions, and (2) multigroup representations in the resonance regions for nuclides having high number densities. In the first case the apparent solution is to perform adjustments on "tailored" sets designed to recognize important cross-section features, and in the second case variables other than the group cross sections must reflect the fact that resonance self-protection factors are independent quantities.

Covariance Representation for Incremental Evaluation

In principle it is possible to perform cross section evaluations including new data by considering an earlier evaluation, together with its variance-covariance

file, as a Bayesian prior.¹⁵ (The same equations are solved in the complete adjustment codes if the new differential data are utilized in a manner similar to the integral data often employed.) This technique has been successfully employed, but with two difficulties of interest here.

Existing covariance evaluations often carefully represent the uncertainties correlated over broad energy ranges but do not represent the uncorrelated uncertainties that exist between cross sections at nearby energies. This approach is often adequate for uncertainty propagation for engineering applications, but for incremental evaluation has the effect that the implied correlations among the prior evaluated cross sections at the energies of new measurements are far too close to unity. The new evaluated cross sections from the code sometimes include apparently unphysical results, results that would be correct if the covariance data supplied to the code had been correct.¹⁶ How may this problem be overcome? It appears that the covariance data required for incremental updating of evaluations is more detailed than the representation needed for propagation of uncertainty in engineering applications. An obvious solution is for the evaluator to derive uncorrelated variance information on a finer grid than that needed by general users of the evaluation. Another approach has been to parameterize the covariance information in terms of correlations effective over small and large energy ranges.¹⁷ This approach is convenient but may not correspond to the actual information content of the underlying experiments and might sometimes correspond to variance-covariance matrices that are not positive definite.

Another problem is that these codes obtain computing economy by assuming that the "new" data being combined with the prior evaluation are uncorrelated with the data base of the prior evaluation. In some cases this assumption is not valid, and then the idea of incremental evaluation is helpful only as a concept.

Final Comments

Reasonably complete representations of the variance-covariance information corresponding to evaluated cross section sets are not available after a decade of experience developing such data for the simple yet important cases. Approaches to more thorough treatments are apparent and further development appears to be warranted for those cross sections most important for practical applications of nuclear data. Until covariance data

10
and uncertainty analysis techniques become somewhat more comprehensive, one cannot be sure about the full extent of the covariance data that will be required in a fully adequate evaluated nuclear data file.

Acknowledgments

The author acknowledges numerous conversations with experimenters and with evaluators of nuclear data active in the Covariance Data Subcommittee of the US Cross Section Evaluation Working Group. Interactions with Francis Perey and Jim Marable has been particularly stimulating.

References

1. As examples: F. Schmittroth, Nucl. Sci. Eng. 72, 19 (1979); D. M. Hetrick and C. Y. Fu, ORNL/TM-7341 (October 1980); W. Poenitz, Argonne National Laboratory, private communication (1982).
2. R. Kinsey, compiler, ENDF-201, BNL-NCS-17541, third edition (ENDF/B-V), 1979; D. W. Muir and R. J. LaBauve, LA-8733-M (March 1981; J. D. Smith and B. L. Broadhead, ORNL/TM-7389 (1981).
3. Many examples are referenced in Sections 5 and 6 of R. W. Peelle and T. W. Burrows, BNL- to be published (1982).
4. F. G. Perey, ORNL/TM-5938 (1979). Reprinted in R. Kinsey, ENDF-102, BNL-NCS-50496 (1979).
5. G. Winkler, D. L. Smith, and J. W. Meadows, Nucl. Sci. Eng. 76, 30 (1980). J. K. Dickens et al. Nucl. Sci. Eng. 74, 106 (1980).
6. L. W. Weston and J. H. Todd, Nucl. Sci. Eng. 79, 184 (1981). D. Larson, ORNL, private communication (1982).
7. F. C. Difilippo, ORNL/TM-5223 (March 1976).
8. R. W. Peelle, "Sensitivity and Uncertainty Analysis of Reactor Performance Parameters," vol. 14 of Advances in Nuclear Science and Technology, Section II-C3, Plenum Press, New York, 1982.
9. J. B. Draft et al. Nucl. Sci. Eng. 62, 119 (1977). B. Strohmaier, S. Tageson, and H. Vonach, Physics Data Nr. 13-2, ISSN 0344-8401 (1980).
10. A. Gandini, CNEN-RP/FI(74)3; E. Greenspan et al. ORNL/RSIC-42, p. 231, 1977; B. Broadhead, Trans. Am. Nucl. Soc. 39, 929 (1981).
11. N. M. Larson and F. G. Perey, ORNL/TM-7485, Nov. 1980.
12. G. de Saussure and R. B. Perez, Annals of Nuclear Energy 9, 79 (1982).
13. J. L. Lucius and J. H. Marable, Trans. Am. Nucl. Soc. 32, 731 (June 1979).

14. D. R. Harris et al. LA-5987 (1975); A. Gandini, CNEN-RT/FI(73)5, (1973); J. B. Draft et al., Nucl Sci. Eng. 62, 119 (1977); J. H. Marable in "Sensitivity and Uncertainty Analysis of Reactor Performance Parameters," vol. 14, of Advances in Nuclear Science and Technology, M. Becker, ed., Plenum Press, New York, 1982; F. G. Perey, ORNL/TM-6267,
15. R. W. Peelle, op. cit., Section IIB. See also Reference 1.
16. D. Hetrick and C. Y. Fu, in LANL informal report T-2-IR(82)-1, E. D. Arthur, ed. (1982).
17. F. Schmittroth and D. W. Wootan, HEDL report TC-1588 (1979).