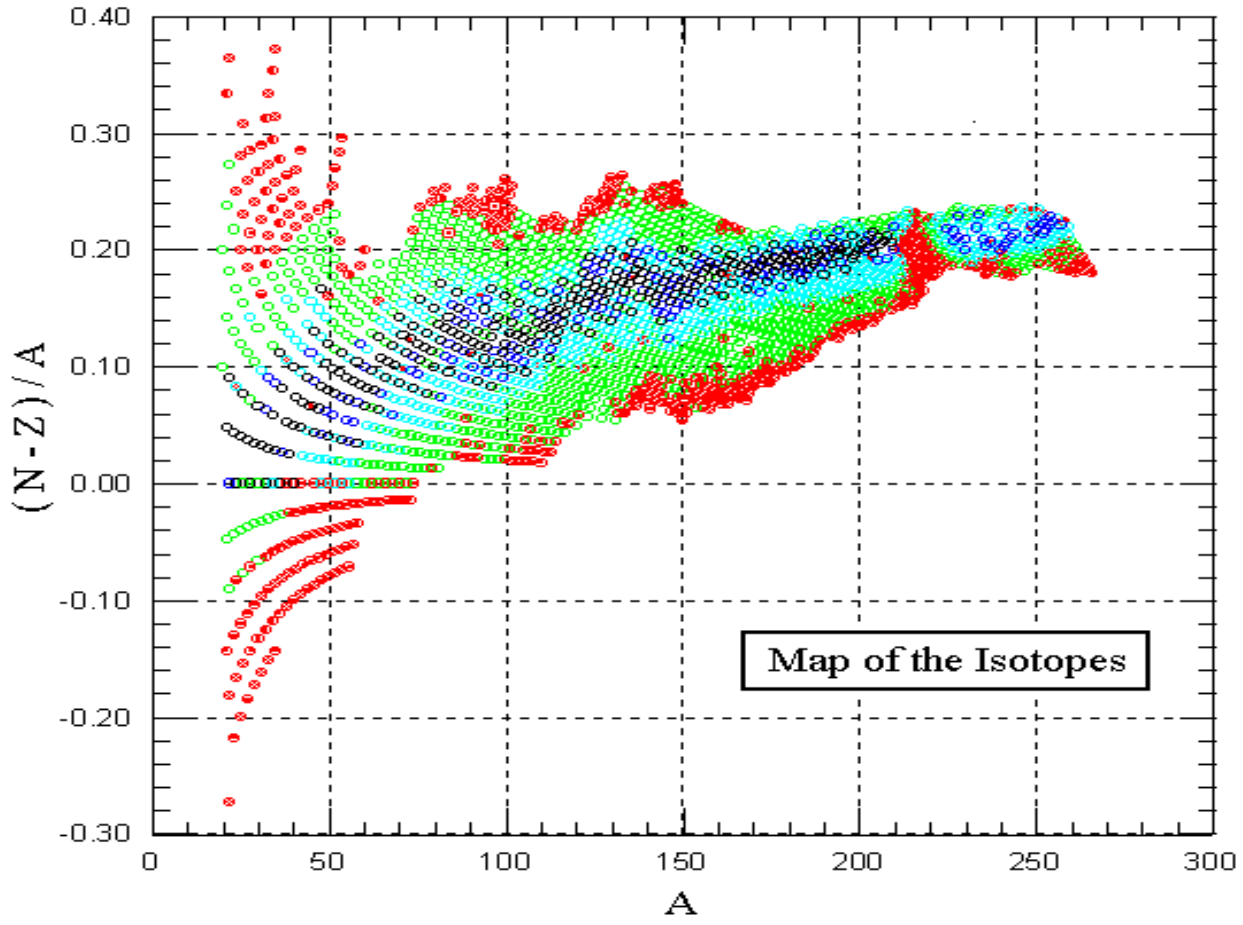
NUCLEAR DATA AND MEASUREMENT SERIES

ANL/NDM-166

A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation

Donald L. Smith

January 2008



NUCLEAR ENGINEERING DIVISION
ARGONNE NATIONAL LABORATORY
9700 SOUTH CASS AVENUE
ARGONNE, ILLINOIS 60439, U.S.A.

About Argonne National Laboratory

Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is in the suburbs of Chicago at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne National Laboratory see http://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 from the 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
Tel: (865)576-8401
Fax: (865)576-5728
Email: reports@adonis.osti.gov

DISCLAIMER

ANL/NDM-166

A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation

Nuclear Data and Measurement Series

by

Donald L. Smith
Nuclear Engineering Division, Argonne National Laboratory

15 January 2008

Office of Science
Department of Energy

UChicago ▶
Argonne LLC

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.

Information Concerning Other Issues of the ANL/NDM Series

Information about all other reports in this Series is available from Argonne National Laboratory at http://www.ne.anl.gov/capabilities/nd/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. If any questions arise concerning the above-mentioned Internet site or the availability of the present report or any other report in the Series, please direct them to:

Dr. Filip G. Kondev
Nuclear Data Program
Nuclear Engineering Division
Building 208/C-162
Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 60439
U.S.A.

Tel: +1(630)252-4484
Fax: +1(630)252-4978
E-mail: kondev@anl.gov

ANL/NDM-166

A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation [a]

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

January 2008

Keywords

NUCLEAR DATA EVALUATION. Monte-Carlo. Probability. Bayes Theorem. Experimental uncertainties. Nuclear modeling uncertainties. Covariances.

---

[a] This work was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Contract DE-AC02-06CH11357.

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

Table of Contents

A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation

Donald L. Smith
Nuclear Engineering Division
Argonne National Laboratory

15 January 2008
(Draft version distributed as an informal NEA-WPEC Memorandum on 15 May 2007)

Abstract

A unified Monte Carlo (UMC) approach to fast neutron cross section data evaluation that incorporates both model-calculated and experimental information is described. The method is based on applications of Bayes Theorem and the Principle of Maximum Entropy as well as on fundamental definitions from probability theory. This report describes the formalism, discusses various practical considerations, and examines a few numerical examples in some detail.

Preface: This document has been prepared in response to a request from the Nuclear Energy Agency (NEA) Working Party on International Evaluation Cooperation (WPEC) to provide support for the activities of Subgroup 24 that are aimed toward development of new methodologies for the evaluation of fast-neutron reaction cross section data.

1. Introduction

The quest for mathematically rigorous nuclear data evaluation methods is motivated by a desire to eliminate – or at least minimize – subjectivity, and it has been enabled by the growing power and sophistication of modern computational resources. The least-squares method – especially the generalized least-squares (GLS) formalism with its capacity to allow inclusion of prior information based on nuclear modeling as well as correlated uncertainties – gained a following in the 1970's (*e.g.*, see refs. [1] and [2]). Some limitations of this approach surfaced in the early 1980's, but it continues to be widely used to this day with various practical, but nevertheless *ad hoc*, "fixes" applied to circumvent these problems (*e.g.*, see ref. [3] which deals specifically with the "Peelle's Pertinent Puzzle" problem, better known as "PPP"). Today, in one form or another, GLS is clearly the most widely employed mathematically objective tool for performing nuclear data evaluations. This report does not attempt to review the various specific methodologies and computer software packages that have been developed and put into use by nuclear data evaluators during the past three decades. Rather, it focuses on describing a single Unified Monte Carlo (UMC) simulation approach that is currently under development. Section 2 describes the basic formalism of this method. Section 3 discusses various practical matters that need to be considered when applying this formalism. Section 4 explores several numerical examples in considerable detail in order

to gain a better understanding of the basic workings of the UMC method. Finally, Section 5 presents some conclusions derived from this investigation.

2. Formalism

The present UMC method, like various others, finds its origins in Bayes Theorem. This theorem is non-controversial and it can be derived easily from the basic postulates of probability theory following some simple steps involving the algebra of probabilities [1]. Bayes Theorem offers a rigorous procedure for learning from experience by defining a simple formula that relates prior and posterior information. For present purposes, we express Bayes theorem in terms of probability density functions rather than actual probabilities. In the following discussion, items expressed in bold font represent vectors and matrices while those in ordinary font are scalars. The symbol "•" will be used for convenience to represent vector (or matrix) multiplication. The symbol "x" signifies scalar multiplication; it is used only in situations where it is needed for clarity.

Let $\mathbf{y}_E$ represent a collection of measured (experimental) quantities with a corresponding covariance matrix $\mathbf{V}_E$ that expresses their uncertainties as well as correlations. Let us suppose that there are "n" elements in the vector $\mathbf{y}_E$ and "n^2" elements in the "n x n" matrix $\mathbf{V}_E$. $\mathbf{V}_E$ must be a symmetric matrix, so the actual number of distinct elements in this matrix is n(n+1)/2. It must also be a positive definite matrix. A "foolproof" method for constructing $\mathbf{V}_E$ is described in Section 3. Furthermore, let $\boldsymbol{\sigma}_C$ represent a collection of quantities calculated from a nuclear model with a corresponding covariance matrix $\mathbf{V}_C$ that expresses their uncertainties as well as correlations. A method for generating this covariance matrix has been reported and discussed earlier by this author, *e.g.* refs. [4] and [5]. We assume that there are "m" calculated quantities and that the corresponding covariance matrix has dimensions "m x m". It must also be symmetric and positive definite. For convenience, we use the symbol "σ" for all the quantities being evaluated even though this collection might include not just cross sections but other observables as well (*e.g.*, angular distributions). The emphasis here is on evaluations derived through combining results from both experiments and nuclear modeling, since this is almost universally the situation encountered by evaluators. In fact, it is the evaluators who generally perform the nuclear modeling exercises. We will not dwell on the issue of how nuclear model parameters and their uncertainties and correlations are chosen to provide the most reasonable values for $\boldsymbol{\sigma}_C$ and $\mathbf{V}_C$. It suffices to say that these results from calculations serve as the "priors" in the ensuing discussion. In other words, we assume that the evaluator begins the evaluation process by generating prior results by means of nuclear modeling and then "refines" the evaluation by incorporating experimental data in the evaluation procedure. If no relevant experimental data exist, then the evaluation will be based on nuclear modeling alone and the evaluator's job is finished. This approach is non-controversial. How could it be otherwise?

In the present context, Bayes theorem is embodied in the following formula:

$$p(\boldsymbol{\sigma}) = C \times \mathcal{L}(\mathbf{y}_E, \mathbf{V}_E \mid \boldsymbol{\sigma}) \times p_0(\boldsymbol{\sigma} \mid \boldsymbol{\sigma}_C, \mathbf{V}_C). \quad (1)$$

In this equation, "p" is the *a posteriori* (posterior solution) probability density function, p_0 is the *a priori* (prior) probability density function, "$\mathfrak{L}$ "is a likelihood function (also a probability density function), and "C" is a normalization constant. This constant is chosen so that the following normalization condition is satisfied:

$$\int_{\mathbb{S}} p(\boldsymbol{\sigma})\, d\boldsymbol{\sigma} = 1 \,, \tag{2}$$

where $d\boldsymbol{\sigma}$ is a volume element (voxel) in the m-dimensional space of possible values for $\boldsymbol{\sigma}$ and $\mathbb{S}$ is the region of that space over which one must integrate in order to achieve convergence. By convergence it is meant that increasing the size of $\mathbb{S}$ would not change the value of the integral in Eq. (2) significantly. We shall see that in practice it is not necessary to know the value of "C" since it cancels in the formulas that are actually used in the Monte Carlo analysis.

It is important to realize that while the components of $\boldsymbol{\sigma}$ are random variable arguments of the indicated functions, $\mathbf{y}_E$, $\mathbf{V}_E$, $\boldsymbol{\sigma}_C$, and $\mathbf{V}_C$ are simply collections of fixed numbers. Since $\boldsymbol{\sigma}$ is a vector, it has the following m components: $\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots, \sigma_m$. The solution to our evaluation problem is completely embodied in the probability density function $p(\boldsymbol{\sigma})$. In probability theory, the "best estimate" value for a random variable, *e.g.*, in this case for σ_i, is defined as its expectation value (better known as "mean value") with respect to the associated probability density function. Therefore,

$$< \sigma_i > = \int_{\mathbb{S}} \sigma_i\, p(\boldsymbol{\sigma})\, d\boldsymbol{\sigma} \quad (i = 1,m) \tag{3}$$

is the evaluated value we seek for this variable.

The same reasoning can be applied to generate a formula for determining elements of the evaluation solution covariance matrix $\mathbf{V}_\sigma$:

$$\mathrm{Cov}(\sigma_i,\sigma_j) = (\mathbf{V}_\sigma)_{ij} = < \sigma_i\, \sigma_j > - < \sigma_i > \mathrm{x} < \sigma_j > \qquad (i,j = 1,m) \,, \tag{4}$$

where <…> represents multivariate integration of the indicated quantities in the same manner as shown for σ_i in Eq. (3). Note that when i = j we obtain the variances from Eq. (4) while the off-diagonal elements (often referred to as "covariances") are obtained when $i \neq j$.

Eqs. (1) through (4) provide us with all that is needed – at least conceptually – to perform an evaluation of the components of $\boldsymbol{\sigma}$ and to determine their covariance matrix $\mathbf{V}_\sigma$. Obviously, we cannot proceed further without addressing several important issues. Here they are:

(i) *Why is the dimension shown here for* $\boldsymbol{\sigma}$ *the same as that for* $\boldsymbol{\sigma}_C$ *, namely, "*m*"?*

The answer is that with a nuclear model one is free to calculate values for the physical parameters of interest at all energies, angles, *etc.*, that one is interested in

representing in the final evaluation. This is certainly not the case for experimental data where myriad factors influence what can be obtained experimentally and what is not obtainable. Therefore, the present methodology (as well as other commonly used ones) generate "prior" results for all the specific energies, angles, *etc.*, that are to be included in the final evaluation. This leads to a one-to-one correspondence between the elements of $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_C$.

(ii) *Are there any conditions concerning the relationship of the prior (model-calculated) and experimental information upon which the UMC evaluation is based?*

Yes, the Bayesian formalism, as embodied in Eq. (1), requires them to be independent, so it is important to make sure that this is the case, at least to a large extent. This point is addressed in some detail in Section 3.

(iii) *What form should the functions "p_0" and "$\mathfrak{L}$" assume?*

Obviously, it is crucial to know exactly what these functions should be since without this knowledge numerical analysis is impossible. Bayes formula, *i.e.*, Eq. (1), offers no specific guidance in this matter. Fortunately, a rigorous solution to this problem comes to us courtesy of the pioneering work on information entropy by Shannon (in the 1940's), Jaynes (in the 1960's), and other statisticians of this period. We will not dwell on the details – they are mentioned in Ref. [1] and other references alluded to therein – but rather will proceed directly to the result. The Principle of Maximum (Information) Entropy tells us that if all we know about a collection of random variables can be summarized by giving their mean values and associated covariance matrix, then the best estimate for the form of the appropriate probability density function is a multivariate normal function (Gaussian). Thus, in our case we have:

$$p_0(\boldsymbol{\sigma} \mid \boldsymbol{\sigma}_C, \mathbf{V}_C) \sim \exp\{-(\tfrac{1}{2})[(\boldsymbol{\sigma} - \boldsymbol{\sigma}_C)^T \bullet \mathbf{V}_C^{-1} \bullet (\boldsymbol{\sigma} - \boldsymbol{\sigma}_C)]\} \ , \tag{5}$$

and by the same reasoning,

$$\mathfrak{L}(\mathbf{y}_E, \mathbf{V}_E \mid \boldsymbol{\sigma}) \sim \exp\{-(\tfrac{1}{2})[(\mathbf{y} - \mathbf{y}_E)^T \bullet \mathbf{V}_E^{-1} \bullet (\mathbf{y} - \mathbf{y}_E)]\} \ . \tag{6}$$

In these formulas $\mathbf{V}_C^{-1}$ and $\mathbf{V}_E^{-1}$ are inverse matrices, "T" denotes the transpose of the indicated vector, and the symbol "~" indicates that the respective normalization constants are not shown explicitly. They are actually not needed as is indicated below. It is now clear why we require $\mathbf{V}_C$ and $\mathbf{V}_E$ to be square, symmetric, positive definite matrices; they have to be inverted. The reason why "$\mathbf{y}$" and "$\mathbf{y}_E$" appear in Eq. (6) rather than "σ"-type variables is that the relationship between the experimental data $\mathbf{y}_E$ and the variables $\boldsymbol{\sigma}$ to be evaluated may be indirect. For example, the experimental data may represent ratios of the variables to be evaluated or they may be integral quantities. In fact, it is appropriate to define $\mathbf{y}$ by the expression $\mathbf{y} = \mathbf{f}(\boldsymbol{\sigma})$, where $\mathbf{f}$ represents a vector collection of m scalar functions $f_1, f_2, \ldots, f_i, \ldots, f_m$ each of whose variables are one or more of the elements of $\boldsymbol{\sigma}$. This point is discussed further in Section 3.

While the conditions that lead to a multivariate normal probability density function for both the prior and likelihood distributions are relatively common ones, it should be noted in passing that other functions may be more appropriate in applications where alternative information is available [1]. For example, if there are estimates of the mean values but no uncertainty information, then an exponential function should be used. Another example might be that both central values and covariance matrices are available but the uncertainties are very large. Under these conditions, lognormal distributions should be used rather than normal distributions [6]. Lastly, if the experimental information is based entirely on raw detectors counts, then a Poisson distribution could be used for the likelihood function.

Combining Eqs. (1), (5), and (6) leads to the expression

$$p(\boldsymbol{\sigma}) \sim \exp\{-(½)\left[\{(\mathbf{y}-\mathbf{y}_E)^T \bullet \mathbf{V}_E^{-1} \bullet (\mathbf{y}-\mathbf{y}_E)\} + \{(\boldsymbol{\sigma}-\boldsymbol{\sigma}_C)^T \bullet \mathbf{V}_C^{-1} \bullet (\boldsymbol{\sigma}-\boldsymbol{\sigma}_C)\}\right]\}\ , \qquad (7)$$

where once again we overlook the normalization constant. It is interesting to note that if we were to apply the assumption that our best solution for the evaluation corresponds to values of the components of $\boldsymbol{\sigma}$ that maximize $p(\boldsymbol{\sigma})$, then we ought to require that

$$[(\mathbf{y}-\mathbf{y}_E)^T \bullet \mathbf{V}_E^{-1} \bullet (\mathbf{y}-\mathbf{y}_E)] + [(\boldsymbol{\sigma}-\boldsymbol{\sigma}_C)^T \bullet \mathbf{V}_C^{-1} \bullet (\boldsymbol{\sigma}-\boldsymbol{\sigma}_C)] = \text{minimum}\ . \qquad (8)$$

This line of reasoning is strictly valid only if $p(\boldsymbol{\sigma})$ is a multivariate normal distribution with respect to the variables $\boldsymbol{\sigma}$. Acceptance of this assumption leads directly to the well-known generalized least-squares formalism (GLS), (*e.g.*, refs. [1] and [2]). However, we will avoid being unduly distracted by this observation in the present development.

Eq. (7) – in combination with Eqs. (2), (3), and (4) - offers a way to carry out the numerical analysis for the proposed UMC evaluation scheme. We note that all the information that we need pertaining to the experimental and model calculated values, as well as their uncertainties and correlations, is embodied in Eq. (7). If we accept Bayes Theorem and the Principle of Maximum entropy, this development of the UMC formalism is quite rigorous up to this point. But, now we are faced with the following challenging problem:

(iv) *Is it necessary to actually calculate the multivariate integrals indicated by the formulas above?*

In principle, the answer to this question is “yes”; in practice the answer is “no”, at least not by the conventional “brute force” approaches used for numerical integration. The reason is that such calculations have been shown to be amenable to analysis by Monte Carlo simulation, at least to precisions which, in principle, are limited only by the number of traced histories.

Let us image pursuing K Monte Carlo histories. For each history we generate a potential solution vector $\boldsymbol{\sigma}_k$ (k = 1,K). Each component of this vector is selected at

random from its associated uniform distribution independently from all the others. A typical sampling range would be defined by:

$$\sigma_{i\text{-}min} \leq \sigma_{ik} \leq \sigma_{i\text{-}max} \qquad (i = 1,m \; ; \; k = 1,K) \quad (9)$$

Expressed another way, σ_{ik} is generated using the following formula:

$$\sigma_{ik} = \sigma_{i\text{-}min} + (\sigma_{i\text{-}max} - \sigma_{i\text{-}min}) \; x \; (RN)_{ik} \qquad (10)$$

where $(RN)_{ik}$ represents a real random number uniformly selected from the interval (0,1). The indicated intervals define a unique "rectangular" region $\mathbb{S}$ in m-dimensional space with volume $\mathcal{V}(\mathbb{S})$ given by the formula

$$\mathcal{V}(\mathbb{S}) = \Pi_{i=1,m} \, (\sigma_{i\text{-}max} - \sigma_{i\text{-}min}) \; . \qquad (11)$$

As mentioned above, the heart of the evaluation process is embodied in Eqs. (2), (3), (4), and (7). Therefore, we proceed now to develop specific forms for these equations that are amenable to Monte Carlo analysis. The equivalent of Eq. (3) is

$$< \sigma_i >_K = [\Sigma_{k=1,K} \, \sigma_{ik} \, p(\boldsymbol{\sigma}_k)] \; / \; [\Sigma_{k=1,K} \, p(\boldsymbol{\sigma}_k)] \; , \qquad (i = 1,m) \qquad (12)$$

while the equivalent to Eq. (4) is

$$\{Cov(\sigma_i,\sigma_j)\}_K = \{(\mathbf{V}_\sigma)_{ij}\}_K = < \sigma_i \, \sigma_j >_K - < \sigma_i >_K \; x \; < \sigma_j >_K \qquad (i,j = 1,m) \; . \qquad (13)$$

To avoid any confusion, we note that

$$< \sigma_i \, \sigma_j >_K = [\Sigma_{k=1,K} \, \sigma_{ik} \, \sigma_{jk} \, p(\boldsymbol{\sigma}_k)] \; / \; [\Sigma_{k=1,K} \, p(\boldsymbol{\sigma}_k)] \; . \qquad (i,j = 1,m) \qquad (14)$$

The sums found in the denominators of Eqs. (12) and (14) are included there to insure proper normalization. The index "K" that appears as subscripts in Eqs. (12) through (14) suggests, and rightfully so, that the values determined using these equations will indeed depend quite strongly on the number of histories K, at least for relatively small K. In fact, for small K the results are most likely meaningless. However, as K becomes large it is anticipated that these quantities should converge toward the values that would be obtain if the corresponding integrations were actually performed as originally indicated in Eqs. (2) through (4). How large does K have to be to achieve acceptable convergence? This can be determined only from experience. In any event, direct deterministic computation of such integrals in those cases where $\boldsymbol{\sigma}$ has very many components is likely to be impractical, even with contemporary computational capabilities. On the other hand, one hopes that a good approximation to these integrals might be obtained through the Monte Carlo simulation route. This approach has been demonstrated to work very well in the analysis of complex nuclear systems, *e.g.*, through widespread applications of codes such as MCNP. Therefore, it seems reasonable to apply in the present context of nuclear data evaluation.

So, the present UMC method essentially amounts to employing Bayes Theorem and the Principle of Maximum Entropy, along with the given prior and measured values and their covariance matrices as constants, to generate a posterior probability density function p for the random variables $\boldsymbol{\sigma}$ that correspond to the evaluation in question. The final evaluated values $< \boldsymbol{\sigma} >$ are therefore first moments (or mean values) of the probability density function p while the elements of the solution covariance matrix $\mathbf{V}_\sigma$ are derived from the second moments of p. The integrals required to determine the mean values and the covariance matrix elements are computed by Monte Carlo integration rather than by deterministic numerical methods. Will this approach work? In principle "yes", but in practice it remains to be demonstrated. Is it a practical approach? Again, this must be ascertained from experience gained in applying the method to realistic situations.

At this point it is worthwhile to enumerate the advantages of the UMC evaluation scheme described above before proceeding to the next section of this Report.

(1) The method is quite rigorous, as long as the errors are not too large, and it involves no conceptual mathematical approximations.

(2) The method employs the Monte Carlo simulation technique fully in contrast to earlier schemes which are either completely deterministic or hybrid combinations of MC and deterministic analyses.

(3) This method does not suffer from limitations arising from the assumptions of linearity which are usually imposed, *e.g.*, in the generalized least-squares (GLS) method. Therefore, it is capable of handling situations where significant non-linearities may result from the structure of nuclear models or from inclusion of complicated experimental data.

(4) The UMC method incorporates the MC simulation concept suggested earlier by this author in the context of generating prior information [4], and extends it to encompass experimental data as well.

(5) Correlated uncertainties can be handled with no obvious limitations.

(6) This method is anticipated to yield results which are consistent with what would be obtained by the widely used generalized-least squares (GLS) method in situations where the uncertainties are small and any encountered non-linear effects are modest.

(7) The method is very straightforward to program for a computer since the formulas are simple. The numerical "bookkeeping" chores should also be relatively modest.

(8) The input values (calculated and experimental) need not be randomly varied in complicated ways in order to deal with correlations, *etc.* All information about the data values, uncertainties, and correlations is completely contained in the fixed arrays of numbers $\mathbf{y}_E$, $\mathbf{V}_E$, $\boldsymbol{\sigma}_C$, and $\mathbf{V}_C$. Furthermore, the two covariance matrices, $\mathbf{V}_E$ and $\mathbf{V}_C$, need to be inverted only once, regardless of the number of histories K considered.

(9) Those variables that DO need to be randomly varied in the UMC approach, *i.e.*, components of the solution vector **σ**, are selected independently by means of uniform sampling within well-defined intervals.

(10) The method can handle a wide variety of experimental data (differential, ratio, integral, *etc.*) as long as well-defined functional relationships exist between the variables that these data represent (**y**) and the solution variables **σ**.

(11) There are no limits, in principle, to the statistical precision that can be achieved for the evaluated solution by the UMC method.

(12) The probability density function $p(\boldsymbol{\sigma})$ does NOT need to be normal with respect to the variables **σ** as is required for the generalized least-squares (GLS) approach to be strictly valid. Yes, Eq. (7) does "appear" to be a normal distribution, but it will not be a true normal distribution with respect to the variables **σ** if the measured quantities (represented by the "y" vectors) involve ratios or other complicated combinations of the components of **σ**.

What is the price to pay for these many advantages? Basically, this method is quite demanding of computational resources. When this author published ref. [1], he "hinted" at this method but dismissed it at the time (1991) as probably impractical because of the computational demands. But, much has changed in the intervening 16 years. Computational power has increased by several orders of magnitude (*e.g.*, even for PC's) while cost has dropped dramatically. Perhaps the UMC method will continue to be viewed as excessively demanding of computational resources, but it should be remembered that the power of computers continues to increase exponentially according to Moore's Law. Also, as this author pointed out at the ND-2004 conference [5], much of the available computational power in research centers still goes unused. There is no reason why individual PC's or workstations should be dormant during nights and weekends. Also, calculations such as these can run in the background while other real-time tasks that are relatively undemanding of processor capability are being performed in the foreground, *e.g.*, word processing. So, it would appear that with the application of some common sense on the part of evaluators concerning the choice of variable sampling limits and procedures, *etc.*, the UMC method should be quite feasible with the computational resources currently available to them. Furthermore, the time and effort required to prepare input information for an evaluation, *e.g.*, insuring that the experimental information which is extracted from compiled libraries such as EXFOR is properly interpreted and prepared for its use in the evaluation, is still likely to exceed the computational time associated with applying the UMC method to produce the evaluation.

3. Practical Considerations

Before proceeding to demonstrate the UMC method through some simple examples, it is necessary to discuss several practical matters.

(i) The Convergence Issue

The UMC method will fail unless it can be demonstrated that the quantities computed by means of MC simulation using Eqs. (12) through (14) actually converge as K becomes large. For this reason, it is important to perform some representative tests of convergence as part of the evaluation process. If Q represents a quantity whose convergence is to be examined, then the general formula used to test convergence is

$$<Q>_K = [\Sigma_{k=1,K}\ Q_k\ g_k] \,/\, [\Sigma_{k=1,K}\ g_k]\ . \qquad (15)$$

Depending on the quantity Q whose convergence is being estimated, g_k might assume the values $g_k = p(\boldsymbol{\sigma}_k)$ or $g_k = 1$. In any event, a simple plot of $<Q>_K$ versus K can be very revealing as a means to estimate qualitatively whether convergence is actually taking place. Several examples of this are shown in Section 4. As a general rule, the quality of convergence will hinge on three considerations: (i) the quality of the random number generator (RNG) used in the MC simulation, (ii) the "volume" $\mathcal{V}(\mathbb{S})$ of the sampling space $\mathbb{S}$, and (iii) the numerical precision employed in the analysis. Very good quality RNG's are available, so this should not be an issue. In choosing the sampling region, one should be certain that it is large enough so that outside this region the magnitude of the posterior probability density function *p* is vanishingly small. More precisely, if a sampled vector $\boldsymbol{\sigma}_k$ is not contained in $\mathbb{S}$, then $p(\boldsymbol{\sigma}_k) \approx 0$. Of course, one could insure this by choosing $\mathbb{S}$ to be very large. However, the penalty to pay for such a conservative choice would be that K would need to be very large to achieve acceptable convergence. Much computational time would then be "wasted" on calculations corresponding to regions of the space $\mathbb{S}$ that yield negligibly to the desired weighted-average quantities. However, if $\mathbb{S}$ is chosen too small then convergence might be observed after a relatively small number of histories K, but this convergence might NOT be to values close to the true values of the integrals which are being approximated. This could be described as an "incomplete integration" effect. Again, experience would have to be the guide in dealing with this issue in each individual instance. Finally, it is certain that a wide dynamic range of real number values will be encountered in computations of the $p(\boldsymbol{\sigma}_k)$ weighting factors. Therefore, a high degree of numerical precision is essential in performing realistic evaluations if one aims to achieve accurate results that are not afflicted by numerical round-off effects. These issues are explored on a small scale in the examples presented in Section 4. They should be more thoroughly investigated for realistic evaluation exercises.

(ii) Compatibility of the Prior and Experimental Information

The input experimental and model-calculated information must be compatible. What do we mean by "compatibility"? In setting up an evaluation exercise, no matter what method is used, the evaluator has to define grid points (or node points if you prefer) that establish the scope of the evaluation. These grid points are characterized by such parameters as incident neutron energy, particle emission angle, *etc.* The final ENDF-formatted files are uniquely defined by these choices. As indicated above, the situation is unambiguous as far as the model-calculated results are concerned. They can be generated in a straightforward manner for all selected node points.

However, for experimental results the situation is murkier. There are two issues involved. Consider the first one. As suggested above, there is a reason why prior and posterior (solution) quantities are labeled "σ" while "y" is used to designate experimental results. The experimental results may be more complicated than simple cross sections. To reinforce this point, and to clarify what is meant by "compatibility", let us consider a particular example. Among the experimental data included in vector $\mathbf{y}_E$, let us consider one particular component, *e.g.*, y_{E7}. We suppose that this experimental value corresponds to a measured differential cross section ratio involving cross sections associated with grid points 6 and 18. Thus, we require that $y_7 = f_7(\boldsymbol{\sigma}) = (\sigma_6/\sigma_{18})$. This must be reflected in the explicit expression for $p(\boldsymbol{\sigma})$. An even more complicated situation arises if the measured quantity is an integral value. Suppose that a particular component of $\mathbf{y}_E$, *e.g.*, y_{E3}, is a spectrum-averaged value measured in a well known spectrum such as the ^{252}Cf spontaneous-fission neutron spectrum. Let the vector $\boldsymbol{\varphi}$ represent this spectrum (normalized) in group format. Then, we require that $y_3 = f_3(\boldsymbol{\sigma}) = \boldsymbol{\varphi}^T \bullet \boldsymbol{\sigma}$, where "T" signifies the transpose of the indicated vector. In practice, matters can become even more complicated. To be perfectly compatible, all input experimental information must be adjusted to correspond to the selected grid points. Thus, *e.g.*, the group representation $\boldsymbol{\varphi}$ for the ^{252}Cf spontaneous-fission neutron spectrum must be expressed in terms of grid-point parameters. The same is true for ratio quantities. A particular example will clarify this point. Referring to the discussion above, let us suppose that the neutron energy corresponding to grid point 6 is 5 MeV while that for grid point 18 is 14 MeV. Then y_7, as defined above, is meant to represent a ratio corresponding exactly to these two energies. However, let us suppose that the measured value y_{E7} actually corresponds to a ratio involving experimental energies 4.9 MeV and 14.1 MeV. Then, it is necessary to adjust the measured value y_{E7} as needed so that it is compatible with y_7. These details are not unique to the present method. In principle, they need to be considered in order to apply correctly any of the more commonly used evaluation techniques, including the generalized least-squares (GLS) method. It easy to see why it is so challenging – and perhaps even folly – to try and automate the evaluation process!

(iii) Generating the Prior Values and Their Covariance Matrix

A MC approach to addressing the problem of uncertainties for priors derived from nuclear modeling was suggested by this author a few years ago [4]. It has been implemented successfully by several laboratories so there is no need to discuss this point in any detail in this report. This approach basically amounts to an error propagation exercise that is performed via MC simulation rather than deterministically. The advantage is that the method does not insist on an assumption of linearity so relatively large uncertainties can be accommodated without sacrificing accuracy. The outcome is unambiguous as long as the number of sampling histories is sufficient. To date, most such analyses have been carried out using 1000 histories. Based on the observed outcomes, this is probably adequate in practice. However, no comprehensive studies of convergence have been carried out, at least to the knowledge of this author. It would be worthwhile to investigate this point and it should not be very demanding to do so.

Current discussions within the nuclear data evaluation community concerning applications of the MC and other fast-neutron cross section data evaluation schemes that rely on nuclear modeling to generate priors revolve around how to best estimate the central values of the nuclear model parameters as well as their uncertainties and correlations. While this is an important issue, these deliberations do not impact on the potential success or failure of the UMC method. All that is needed to proceed with the present approach is that $\boldsymbol{\sigma}_C$ and $\mathbf{V}_C$ be provided, since they are required to construct the prior probability density function p_0. This prior information, and the experimental information that is to be merged with it to generate an evaluation, should in principle be independent. Therefore, it is very desirable, for conceptual reasons, that the selection of nuclear model parameters and their uncertainties be influenced as little as possible by the specific nucleus for which the evaluation in question is being carried out. A reasonable way to achieve an adequate degree of independence is for the choice of nuclear model parameters used to generate the prior to be guided by global considerations, *e.g.*, by knowledge gained from consideration of a wide range of nuclei across the Periodic Table rather than strictly by narrow regional or local nuclear model behavior. It is a matter of judgment and experience as to how this should be approached in practice. As long as an evaluator is aware of this issue, and takes some precautions aimed at achieving a decent measure of independence for the prior information and the experimental information for a specific nucleus, it is anticipated that reasonable evaluated results can be produced. Experience shows that the correlations for the prior values tend to be rather large. It appears that the nuclear models themselves, rather than the specific properties of the model parameters, are the major sources of correlations encountered when priors are determined from nuclear modeling.

(iv) Preparation of the Experimental Data

The need to adjust experimental data so that they will correspond to calculated values at the selected grid points has been mentioned above. Here we are concerned with the actual quality of the compiled experimental data. There can be no disputing the fact that utilization of poor quality experimental data and incomplete or improperly constructed covariance matrices can thwart the evaluation process. The rule "garbage in, garbage out" applies regardless of the evaluation method. It is no less true in applying the present UMC method. The need for weeding out bad data, applying adjustments for changes in standards, enhancing some unrealistically small assigned uncertainties, and other data "preparation" steps is widely acknowledged by evaluators as necessary if one is to achieve reasonable evaluated results. This is reflected in the recent decision by WPEC to establish Subgroup 30 with the objective of eliminating some of the worst defects currently lurking in the database EXFOR. There can be little doubt that attending to the sorry status of the compiled experimental database – a resource that represents many decades of effort and huge expenditures of financial resources – is one of the major challenges now facing the nuclear data evaluation community.

In this report we outline the procedure for preparing the covariance matrix for a single experimental data set. This procedure is discussed extensively in ref. [1] so the discussion in this report is brief to avoid unneeded repetition. Let us suppose that we have

a collection of N experimental data values. Furthermore, assume that there are B distinct (independent) sources of experimental error ($\beta = 1,B$). Then, assume that $e_{i\beta}$ ($i = 1,N$; $\beta = 1,B$) is the magnitude of the absolute error component for the i^{th} data point corresponding to the β^{th} error source. Finally, assume that $C_{\beta ij}$ ($i,j = 1,N$; $\beta = 1,B$) is an element of the matrix $\mathbf{C}_\beta$ that describes the correlations associated with the β^{th} error source. The $(i,j)^{th}$ element of the covariance matrix for the entire experimental data set is then given by the formula

$$V_{ij} = \Sigma_{\beta=1,B} \ (e_{i\beta} \times C_{\beta ij} \times e_{j\beta}) \ . \qquad (16)$$

If the β^{th} error component is completely uncorrelated (random) for all data points in the set, then $C_{\beta ij} = \delta_{ij}$ (the Kronecker delta function that equals 1 when $i = j$ and 0 when $i \neq j$). When a particular error component is 100% correlated between all the data points, then $C_{\beta ij} = 1$ for all i and j. Partial correlations (*i.e.*, neither 0 nor 1, and possibly negative) are also possible for certain types of errors, but their estimation tends to be more subjective.

While this procedure is reasonably straightforward in theory, what is an evaluator actually to do in practice? Very few authors of experimental data sets provide complete covariance matrices for their reported results, and many fail even to identify the distinct error components involved in their experiments. An evaluator could choose to assume that all the errors are uncorrelated. However, it has been shown that the danger of doing this in situations where extensive experimental data are available is that the final uncertainties for the evaluated results will be unrealistically small. It is well known that a neglect of correlations generally leads to excessively small evaluated uncertainties. The experience of this author has been to handle this dilemma as follows: First, make sure that the given total errors are at least realistic. If not, an evaluator may wish (using caution) to enhance the given errors to values more in line with what one might reasonably expect them to be for an experiment of this nature. Concerning correlations, the evaluator should keep in mind that total error correlations (positive or negative) with magnitudes in the range < 0.3 should be treated as weak while those > 0.7 (but of course ≤ 1) should be considered as strong. All correlations in the mid range of (0.3 to 0.7) are moderate. As an example, if the total error is 5% and the correlated error is 4%, then the random error will be 3% and the correlation will be 0.8 (which is relatively strong). If an author states only the total error and provides no further details that would enable an evaluator to assign objective correlations, then the evaluator might decide to apportion the total uncertainty evenly between random and fully correlated components (*e.g.*, 5% random and 5% fully correlated yields $\approx$ 7% total error). Then the correlation coefficient would be 0.5 which is moderate. This is a reasonable choice under such circumstances. The evaluator could defend this choice by referring to a well known philosophical principle known as *Occam's Razor*, attributed to William of Occum (1320): *"When confronted with multiple options for addressing a problem, choose the simplest one."*

(v) Consistency of the Model-Calculated and Experimental Information

We need to distinguish "consistency" from "compatibility". The issue of "compatibility" is discussed above. By examining the data "consistency", we are

studying the relative scatter of the results that are to be used in an evaluation. In the generalized least-squares formalism (GLS) it is shown that there exists a chi-square (χ^2) test of input data consistency that can be applied before the GLS analysis is performed (*e.g.*, refs. [1] and [2]). However, since this test involves only the input data it seems reasonable to consider applying it in the present UMC methodology. The formula used in this test of consistency is as follows:

$$\chi^2/(\text{d.o.f.}) = [(\mathbf{y}_E - \mathbf{q})^T \bullet (\mathbf{V}_q + \mathbf{V}_E)^{-1} \bullet (\mathbf{y}_E - \mathbf{q})] / n \, . \tag{17}$$

This expression, as well as some of the quantities appearing therein, requires some explanation. In the GLS formalism, the degrees-of-freedom parameter (d.o.f.) is just the number of experimental data values "n". The quantities $\mathbf{y}_E$ and $\mathbf{V}_E$ require no explanation; they are defined above. The vector $\mathbf{q}$ is the collection of n calculated equivalents to the measured data based on prior values of the variables to be evaluated (NOT on solution values), *i.e.*, on $\boldsymbol{\sigma}_C$ and $\mathbf{V}_C$. In other words, $\mathbf{q} = \mathbf{f}(\boldsymbol{\sigma}_C)$. Furthermore, $\mathbf{V}_q$ is an "n x n" covariance matrix which is computed by propagating the errors of $\boldsymbol{\sigma}_C$, as reflected in the covariance matrix $\mathbf{V}_C$, through to $\mathbf{V}_q$ via the functional relationships represented by $\mathbf{f}$. It is clear from Eq. (17) that the matrix $(\mathbf{V}_q + \mathbf{V}_E)$ needs to be inverted, so it must first be tested for positive definiteness. Basically, Eq. (17) provides a means to compute the scatter of the experimental data relative to equivalent calculated values (the word "equivalent" is significant here), scaled by the combined uncertainties of the experimental and calculated results. The general rule to follow is that when $\chi^2/(\text{d.o.f.}) \leq 1$, then the uncertainties in the evaluated results generated by the UMC method ought to be accepted as they are. However, if $\chi^2/(\text{d.o.f.})$ is significantly larger than unity, one might consider enhancing all the evaluation solution uncertainties by the factor $[\chi^2/(\text{d.o.f.})]^{½}$ without altering the correlations. Taking this step amounts an to an honest admission by the evaluator that the actual source (or sources) of discrepancies in the input information cannot be uncovered so he (or she) chooses to "spread the guilt" uniformly by enhancing all the solution uncertainties. Perhaps this approach should be viewed as resorting once again to *Occam's Razor*.

4. Numerical Examples

We begin this section by stating that "… the proof of the pudding is in the eating". With this in mind, we consider a few relatively simple examples that seem appropriate to the development of the UMC method, at least from a numerical perspective. For this purpose, EXCEL spreadsheet routines were used in all the calculations except for the one pertaining to GLS. The idea behind this choice is that if the present method can be demonstrated to be viable by using such a relatively unsophisticated computational tool, then so much the better. For the most part, 1000 histories have been employed for each case considered. However, in a few instances 10 repetitions of 1000 histories each were undertaken. These results were then further averaged to yield what amounts to the equivalent of results obtained from 10,000 histories (to improve statistical precision).

Example 1

In this example we consider a simple test of "randomness" for a table of 1000 random numbers (RN) generated by EXCEL. If a collection of numbers drawn uniformly from the range (0,1) is truly random, then we would expect their average <RN> to be precisely 0.5. Any deviation from this outcome provides some measure of the departure from randomness or, more likely, a reflection of statistical uncertainty associated with the limited sample size. If $(RN)_k$ is one member of this collection, then our test of convergence to the theoretical value is based on the formula

$$<RN>_K = [\Sigma_{k=1,K} (RN)_k]/K \,, \qquad (18)$$

which is just a special case of Eq. (15) with all $g_k = 1$. In this example, the result obtained for K = 1000 is 0.4906 (rounded to four significant figures). This is not unreasonable considering that a 3% uncertainty might be anticipated from statistical considerations. Fig. 1 is a plot of the actual values of the 1000 randomly selected numbers.

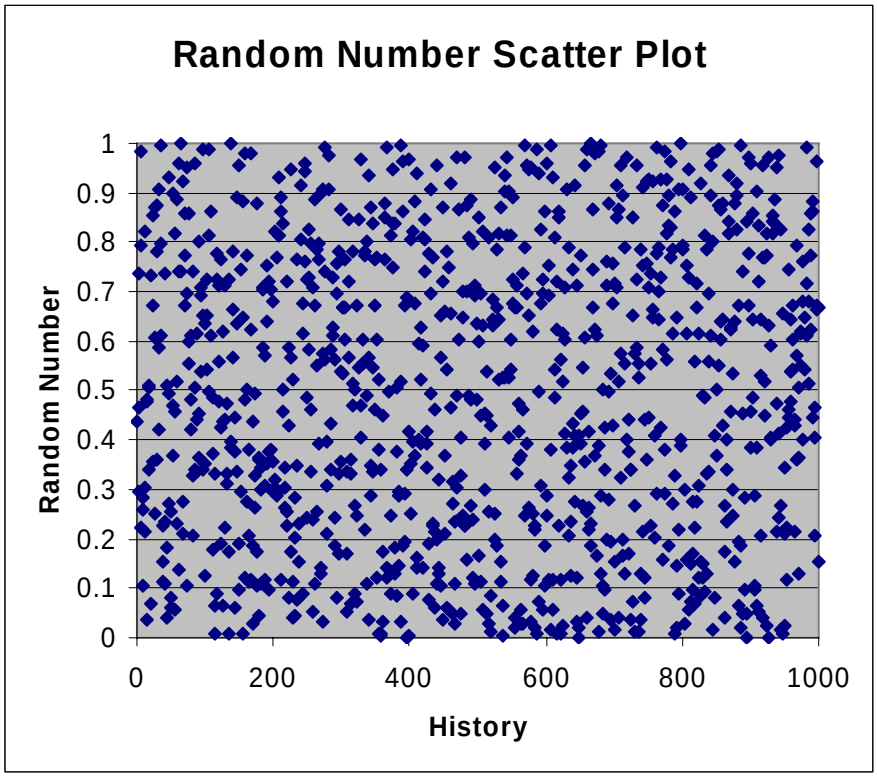


Figure 1: Scatter plot showing 1000 uniformly selected random numbers from the interval (0,1)

Is there a pattern to be discerned in Fig. 1? Humans have a notorious reputation for seeing patterns where none really exist so it's up to the reader to decide the answer to this question for himself (or herself). What is even more interesting is to plot the results

of 1000 applications of Eq. (18), *i.e.*, for K = 1 to 1000, based on the random number set generated for this example. The results appear in Fig. 2.

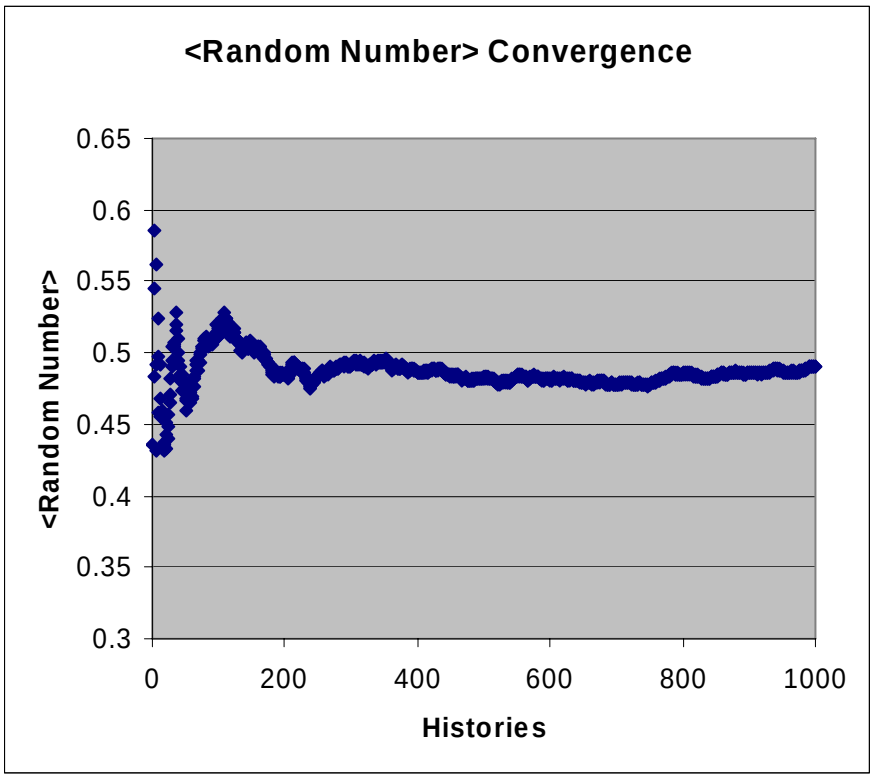


Figure 2: Plot of $\langle RN \rangle_K$ as a function of K for the random number set of Example 1

In this case it is clear that a pattern does emerge. As K becomes large, the values of $\langle RN \rangle_K$ do appear to approach 0.5. However, it is also clear that the "goal" has not been quite reached by K = 1000, and that for many values of K < 1000 the computed average is indeed significantly different from the theoretically "correct" value of 0.5.

This simple exercise has little to do directly with the UMC method under discussion in this report. It is included here to offer the reader some insight as to what numerical issues need to be considered when applying the UMC method and to illustrate the concept of "convergence".

Example 2

Let us proceed to a slightly more complicated situation. Instead of one set of 1000 random numbers we generate two such sets by selecting all the individual numbers uniformly from the interval (0,1). Let us denote the individual random values from these two sets by x_k and y_k, respectively. Next, we generate 1000 values z_k by computing products of x_k and y_k. Thus, $z_k = x_k \times y_k$. A scatter plot of the 1000 values of z_k is shown in Fig. 3.

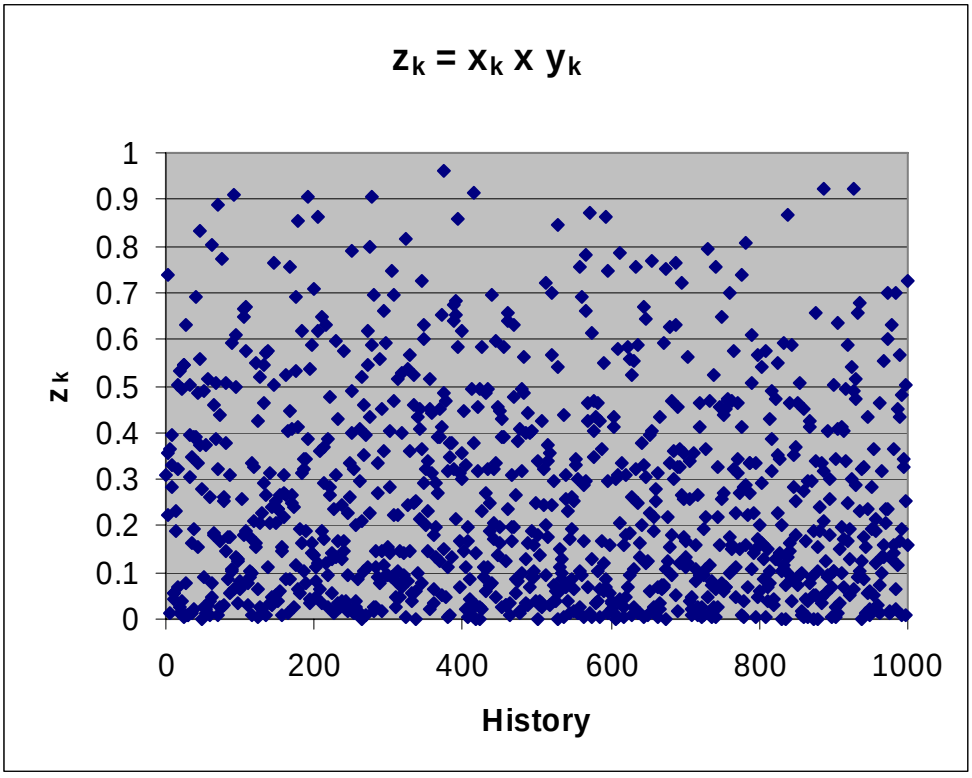


Figure 3: Scatter plot of random values $z_k = x_k \times y_k$ where x_k and y_k are independent random numbers uniformly selected from the interval (0,1)

The reader will find no difficulty in discerning a pattern in Fig. 3. There is a decided "clumping" of z_k values toward the lower end of the interval (0,1). Why? Even though x_k and y_k are random and uniformly distributed in this interval, their product is a non-linear function of these two variables so it will NOT be uniformly distributed in (0,1). Such is the nature of non-linear behavior.

Next, we look for convergence of $\langle z \rangle_K$ as K approaches 1000 in much the same manner as was done in Example 1. The formula that we plot is

$$\langle z \rangle_K = [\Sigma_{k=1,K}\, z_k]/K \, . \qquad (19)$$

The results are shown in Fig. 4.

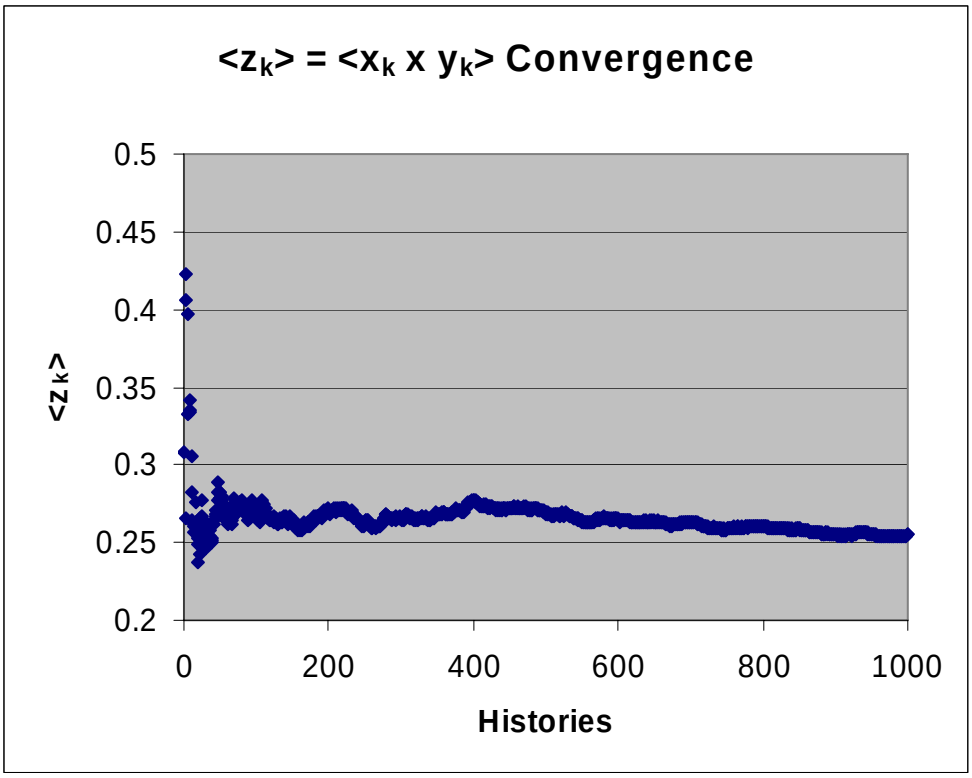


Figure 4: Plot of $\langle z \rangle_K$ as a function of K based on two sets of uniformly distributed random numbers drawn from the interval (0,1)

The behavior shown in Fig. 4 is qualitatively similar to what is observed in Fig. 2 of Example 1. Convergence toward the theoretical value of 0.25 is observed but it is slow. The value of $\langle z \rangle_K$ observed for K = 1000 is 0.2549 (rounded to four significant figures). The reader will also note that 0.25 is the value of the simple double integral defined explicitly by the expression $\int_0^1 \int_0^1$ (xy) dxdy. Thus, we have demonstrated the computation of a multi-variable integral by Monte Carlo simulation through this example.

Example 3

Next we consider an example which in its design is a step closer to what we actually will need to consider if we are to demonstrate the UMC method. Suppose that x and y are independent random variables with the following properties: $\langle x \rangle = 0.5$, SD(x) = 0.2, $\langle y \rangle = 0.5$, SD(x) = 0.2. Here, $\langle \ldots \rangle$ signifies mean value and "SD" stands for "standard deviation". Thus each variable has a mean value of 0.5 and standard deviation of 0.2. This represents quite a large error (40%) for each variable. As a consequence, these variables probably ought to be represented by lognormal probability distributions [6]. Nevertheless, for present purposes we shall consider them to be normal and, furthermore, we will sample these variables uniformly in the range (0,1). This essentially limits the sampling space to a little more than two standard deviations (2-Sigma) to each side of the respective mean values. It will be interesting to learn whether this is sufficient.

By definition, the bivariate probability density function governing the two random variables x and y is given by

$$p(x,y) = C \times \exp[-(\tfrac{1}{2}) \times (1/0.04)(x^2 + y^2)] = C \times \exp[-(12.5) \times (x^2 + y^2)] \,, \qquad (20)$$

where "C" is a normalization constant that we assume to be unity for convenience.

Application of the present methodology involves Monte Carlo calculations based on Eqs. (7), (9), (10), (12) through (14), and (15). The number of histories traced in this exercise is 1000. The first test is to examine the convergence of $\langle p(x,y)\rangle_K$ with increasing K. This is accomplished using Eq. (15) with $Q_k = p(x_k,y_k)$ and $g_k = 1$. The results are plotted in Fig. 5.

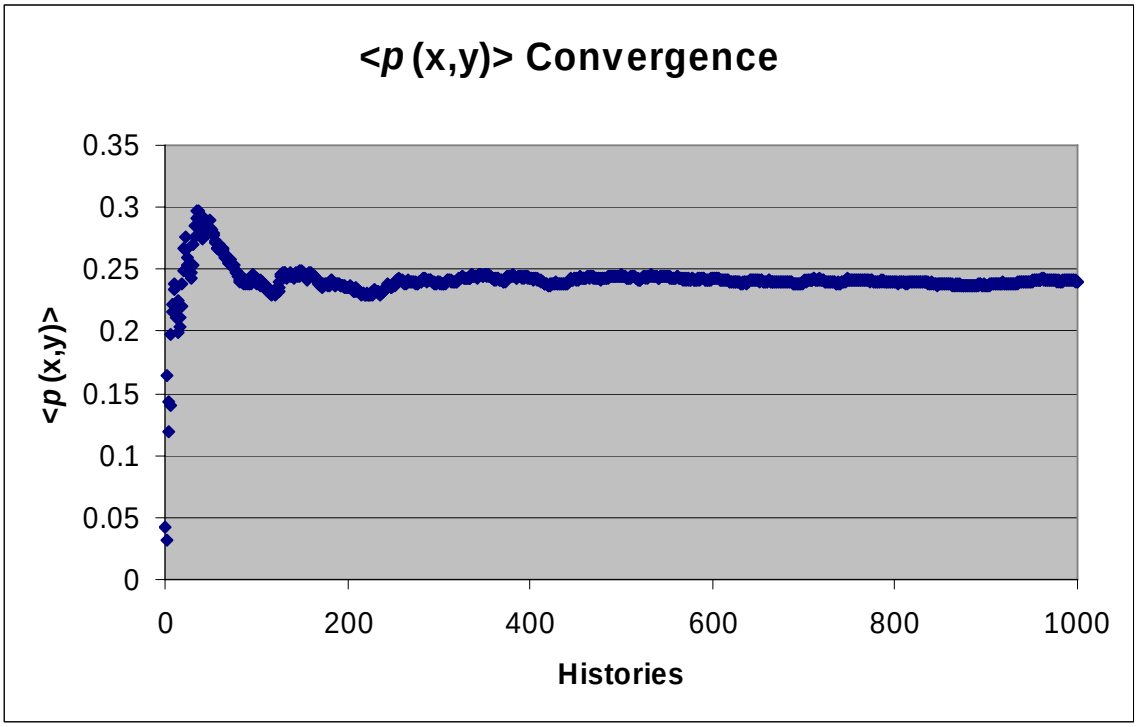


Figure 5: Convergence of $\langle p(x,y)\rangle_K$ versus K based on Eq. (20)

It is evident that reasonably good convergence is achieved as K approaches 1000. Next, we examine the convergence of $\langle x\rangle_K$ using Eq. (15) with $Q_k = x_k$ and $g_k = p(x_k,y_k)$. The results are plotted in Fig. 6.

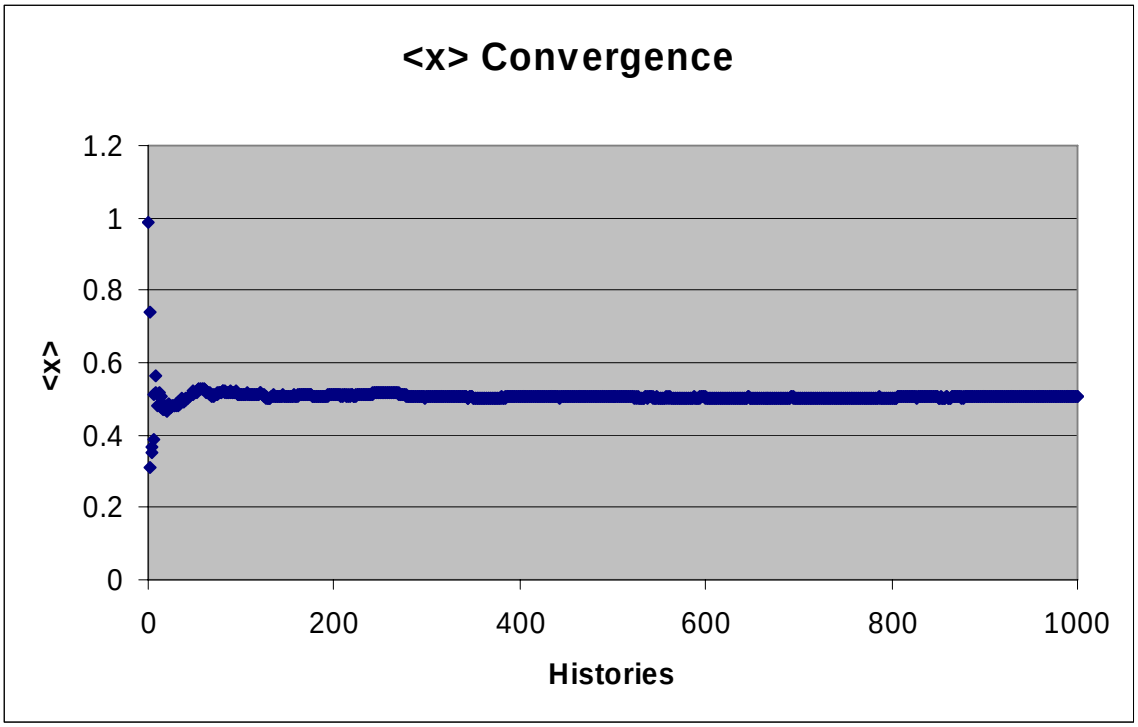


Figure 6: Convergence of $\langle x \rangle_K$ versus K based on Eq. (20)

Clearly, the convergence of $\langle x \rangle_K$ with increasing K is quite rapid. In fact, very little change is seen after ≈ 300 histories. The complete results from this exercise are summarized in Table 1.

Table 1: Summary of results obtained for Example 3

Quantity	MC	Actual
<x> =	0.505466349	0.5
var(x) =	0.036509305	0.04
SD(x) =	0.191074082	0.2
<y> =	0.499706076	0.5
var(y) =	0.038696771	0.04
SD(y) =	0.19671495	0.2
Cov(x,y) =	0.000679885	0

Note: The values in this table were copied directly from the EXCEL worksheet

The results obtained from the present MC exercise are certainly consistent, within statistically reasonable expectations, with the actual values that were known at the outset of this exercise.

Example 4

The last example to be considered here includes all the essential aspects of a realistic evaluation problem, but it is simplified to the point where the analysis is both straightforward and transparent. In spite of the simplicity, it serves the purpose to demonstrate the viability of the UMC method proposed in this report.

The mathematical "model" chosen for this example is defined by the function

$$s(E) = p_1 \times E \times \exp(-E/p_2) \,, \qquad (21)$$

where p_1 and p_2 are the "parameters" of the model and E is a continuous variable. We arbitrarily choose the mean values and errors (standard deviations) of these model parameters as follows: $\langle p_1 \rangle = 10$, $SD(p_1) = 1$, $\langle p_2 \rangle = 2$, and $SD(p_2) = 0.2$. Furthermore, we assume these parameter errors to be uncorrelated. Our assumption indicates that the parameter uncertainties are 10%. A plot of this function is shown in Fig. 7.

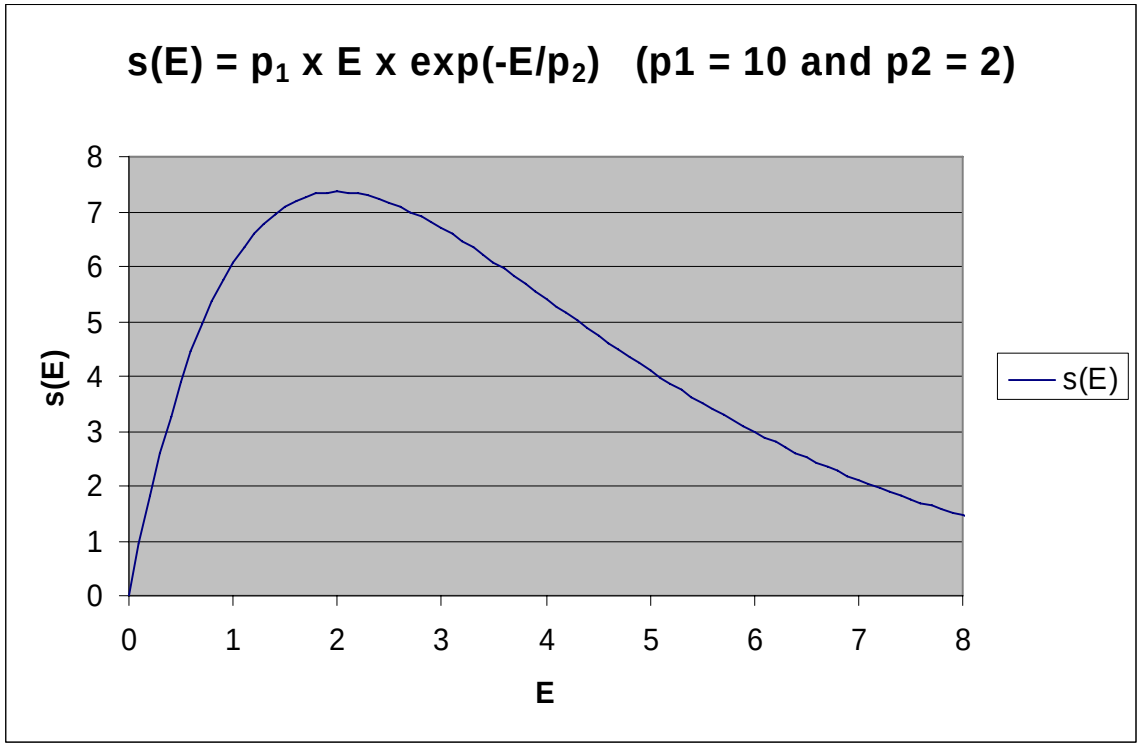


Figure 7: Plot of $s(E) = p_1 \times E \times \exp(-E/p_2)$ with parameters $p_1 = 10$ and $p_2 = 2$

Although it is irrelevant for the present exercise, the shape shown in Fig. 7 is not radically different from what one might expect for the differential cross section of a low-threshold neutron reaction. For example, compare the shape in Fig. 7 to the shape of the $^{58}Ni(n,p)^{58}Co$ differential cross section, as shown in Fig. 8.

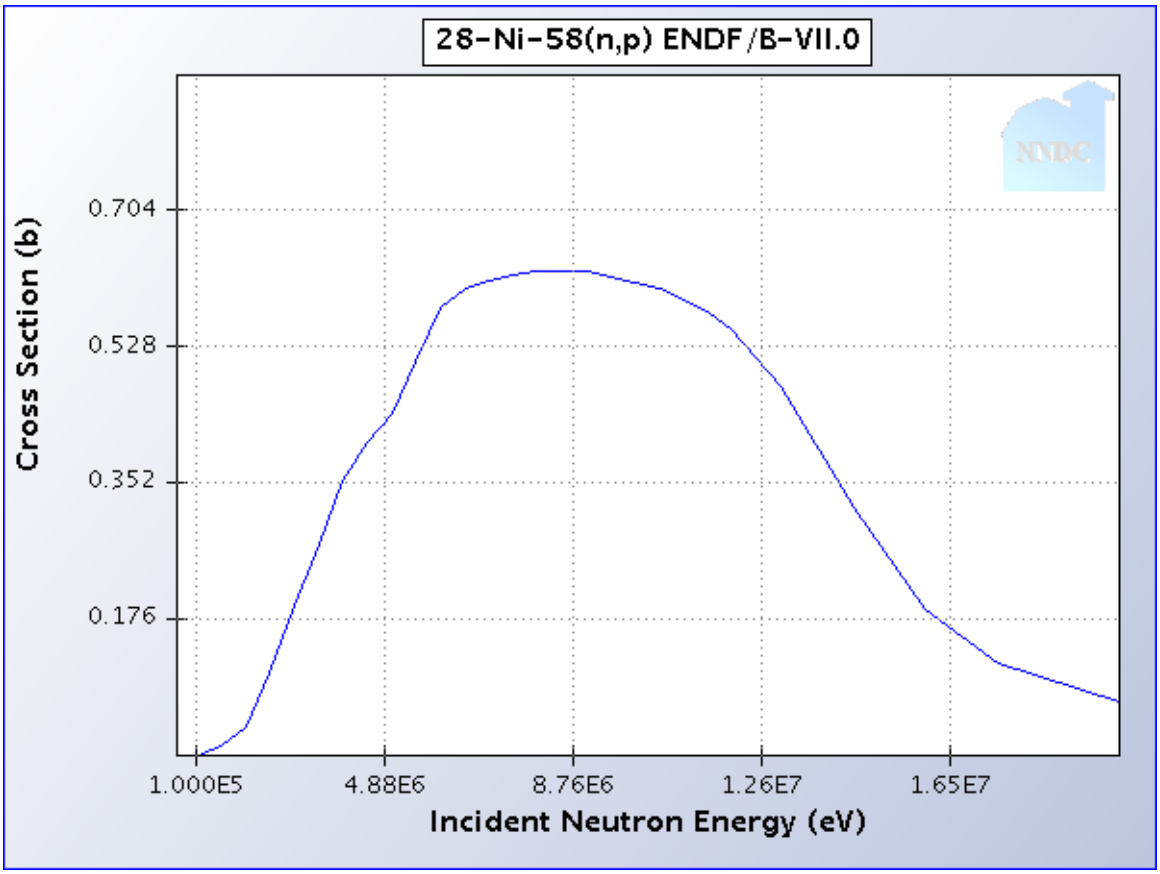


Figure 8: Plot of the ^{58}Ni(n,p)^{58}Co cross section retrieved from BNL-NNDC

To narrow the scope of this example, we choose to consider only two grid points for the present evaluation. These are defined by E = 1 and E = 3 (units are irrelevant here). Using the "model" described above, it is straightforward to determine the two calculated values which form a vector that we label $\mathbf{s}_C$. The two components are s_{C1} and s_{C2}. To insure that there will be no confusion regarding this exercise, we have replaced "σ" by "s" (for both model-calculated and experimental results) since we are not dealing with actual cross sections in this example. Continuing with the analysis, we use the MC methodology described in ref. [4] to determine the corresponding covariance matrix $\mathbf{V}_C$. The parameters p_1 and p_2 are sampled randomly from uniform distributions whose limits are consistent with the assigned mean values and standard deviations. The details are omitted for brevity. Values for the priors to be used in Example 4 are given in Table 2.

Table 2: Calculated values that constitute the two elements of $\mathbf{s}_C$ along with their corresponding covariance matrix $\mathbf{V}_C$; this analysis is based on the two-parameter model defined by Eq. (21) with the parameter values and their uncertainties as indicated above.

Grid Index (i)	E_i	s_{Ci}	SD(s_{Ci})	Error (i)	Covariance Matrix ($\mathbf{V}_C$)	
1	1	6.0653	0.67896	11.19 %	0.46099	0.71303
2	3	6.6939	1.2106	18.09 %	0.71303	1.46556

Note: the numerical values shown here are rounded to better fit the table, but all numerical calculations (including the MC simulation) have been performed using the full precision inherent to EXCEL

The covariance matrix $\mathbf{V}_C$ is positive definite so it can be inverted. The correlation matrix for the calculated values of $\mathbf{s}_C$ is 0.8675 (to four significant figures), or 86.75%. This strong correlation is typical of what one generally obtains from calculations using much more complicated and realistic nuclear models. Incidentally, a deterministic "hand" calculation of the elements of $\mathbf{V}_C$ was also carried out. These results agree extremely well (to several significant figures) with the MC results in Table 2. This supports the MC approach described in ref. [4]. So, calculated values intended to serve as the priors for the present example, along with their covariance matrix, are produced using the simple model described above. This yields half of the input information required for this UMC evaluation exercise.

The next step is to introduce two pseudo "experimental" data points, represented by $\mathbf{s}_E$, which are intended to be used in conjunction with the calculated prior values in the present UMC evaluation. We choose values for these "experimental" data that are not in very good agreement (inconsistent) with the corresponding model-calculated ones. These "experimental" data, however, are entirely comparable to the model-calculated results (no ratios, no departures from grid points, *etc.*). In other words the two "experimental" data points, s_{E1} and s_{E2}, correspond to E = 1 and 3, respectively, so they can be compared directly with their calculated counterparts, s_{C1} and s_{C2}. Thus, these "experimental" results can be used directly in the evaluation process with no additional preparation required. We stipulate that these "experimental" results are fairly accurate (5% uncertainty for each data point) and that their uncertainties are strongly correlated (to the extent of 80%). This is reflected in the data covariance matrix $\mathbf{V}_E$. Table 3 summarizes the "experimental" input information for Example 4.

Table 3: "Experimental" values that form the two elements of $\mathbf{s}_E$ along with their corresponding covariance matrix $\mathbf{V}_E$ for use in Example 4

Grid Index (i)	E_i	s_{Ei}	SD(s_{Ei})	Error (i)	Covariance Matrix ($\mathbf{V}_E$)	
1	1	6.5	0.325	5 %	0.105625	0.0793
2	3	6.1	0.305	5 %	0.0793	0.093025

The calculated and "experimental" values (minus error bars) are plotted in Fig. 9.

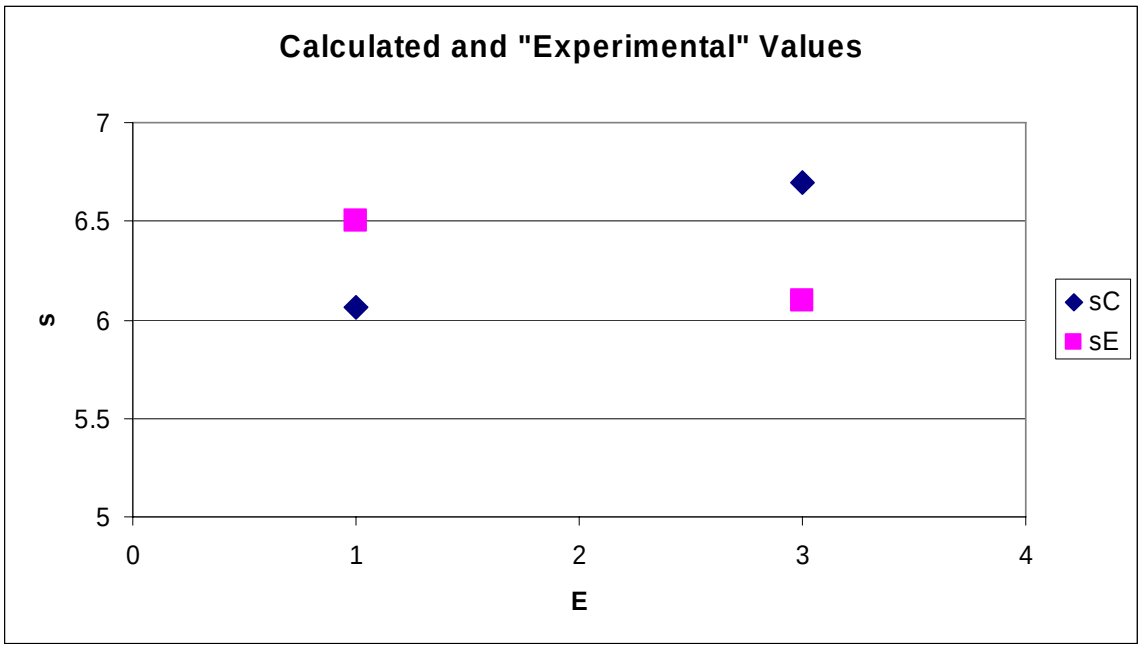


Figure 9: Plot of calculated and "experimental" data considered in Example 4

This exercise of merging strongly correlated calculated results having relatively large uncertainties with "experimental" values having much smaller uncertainties, but also strong correlations, offers a worthy challenge to test the present UMC evaluation method. Our goal is to determine by the UMC approach the evaluated quantities $\mathbf{s}$ and $\mathbf{V}_s$, based on the given input data $\mathbf{s}_E$, $\mathbf{V}_E$, $\mathbf{s}_C$, and $\mathbf{V}_C$. Before beginning the UMC analysis, we attend to two more chores. First, it is necessary to decide on the limits that define the sampling space. As noted earlier, the range must be large enough to ensure adequate sampling of the probability distribution $p(\mathbf{s})$ in this example, but not so large that most of the computational time is spent in regions of low probability. To start off, we choose to randomly sample s_1 uniformly in the interval (4,9) and s_2 uniformly in the interval (2,10). Remember that there are no correlations to consider in carrying out this sampling. These intervals are generously large considering the input data as well as solution results we are likely to obtain. In fact, we choose to refer to this as the "Wide Limits" condition. It's better to operate under relaxed conditions at the outset and then tighten the constraints later if appropriate. The second task is to perform the evaluation using the well-established generalized least-squares algorithm (GLS), with $\mathbf{s}_E$, $\mathbf{V}_E$, $\mathbf{s}_C$, and $\mathbf{V}_C$ as input, in order to establish a baseline solution for comparison purposes. This second task was carried out using code GLSMOD, as described in ref. [2].

A single UMC simulation exercise was performed first using K = 1000 histories, the formalism described in Section 2, the "experimental" and calculated input values, and the "Wide Limits" condition. We do not repeat the formulas or procedures here since they have already been mentioned several times above. In particular, we avoid giving a specific expression for the probability density function $p(\mathbf{s}) = p(s_1,s_2)$ since its form should by now be obvious to the reader from the preceding discussions. Following this

first exercise, the convergence of $\langle p(s_1,s_2)\rangle_K$ as a function of K was examined. The results are shown in Fig. 10. We assume C = 1 for convenience.

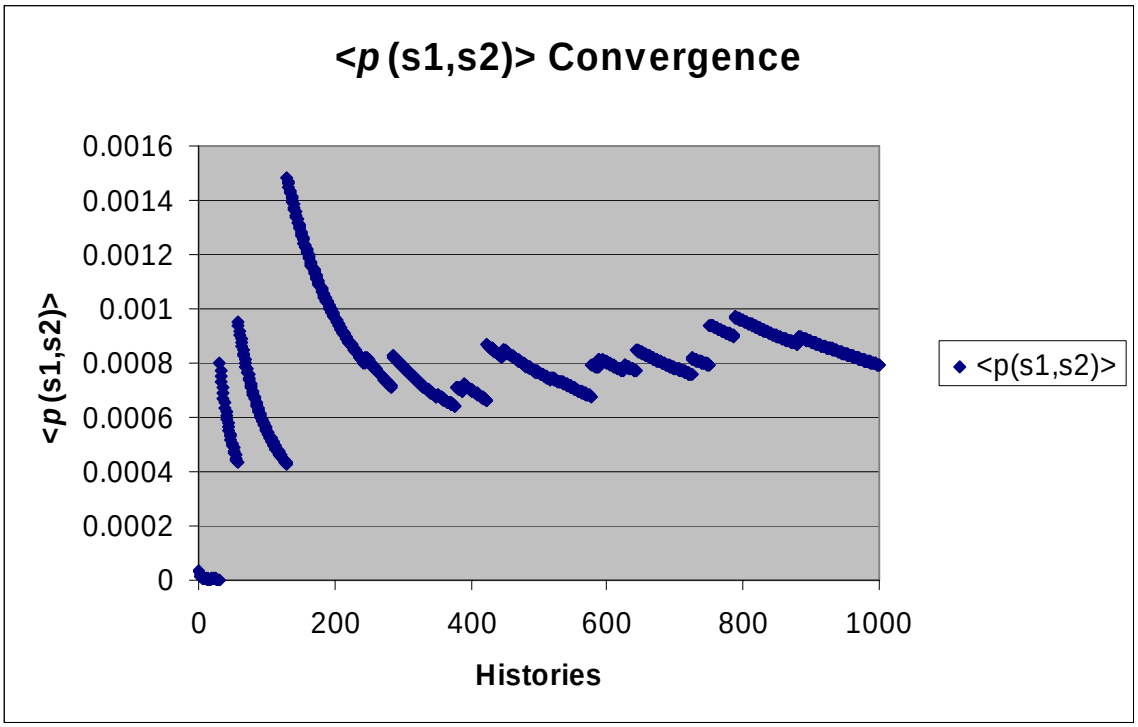


Figure 10: Plot of $\langle p(s_1,s_2)\rangle_K$ as a function of K for Example 4

The profile in Fig. 10 looks like an approaching tsunami! It is not surprising that such a dramatic pattern would be generated since the probability density function $p(s_1,s_2)$ varies over many, many orders of magnitude as the variables s_1 and s_2 are sampled uniformly throughout the space of values allowed by the "Wide Limits" sampling rule. In fact, different strings of 1000 randomly selected variables s_1 and s_2 generate quite different patterns. What matters, though, is whether there is clear evidence of convergence of $\langle p(s_1,s_2)\rangle_K$ for values of K approaching 1000. An examination of Fig. 10 would suggest that no such clear evidence exists. Therefore, K = 1000 histories is probably insufficient, at least when the sampling limits are set rather wide as is the case for the "Wide Limits" condition. With this in mind, nine additional independent sets of 1000 randomly selected pairs of values s_1 and s_2 were generated and the same analysis as described above was carried out for each set. The accumulated results were then averaged to yield what we would anticipate achieving with 10,000 histories. Figure 11 is a rather busy plot that shows the "experimental" and calculated values, the results of the GLS evaluation exercise, the results from 10 independent UMC evaluations with 1000 histories each and, finally, the averaged UMC results.

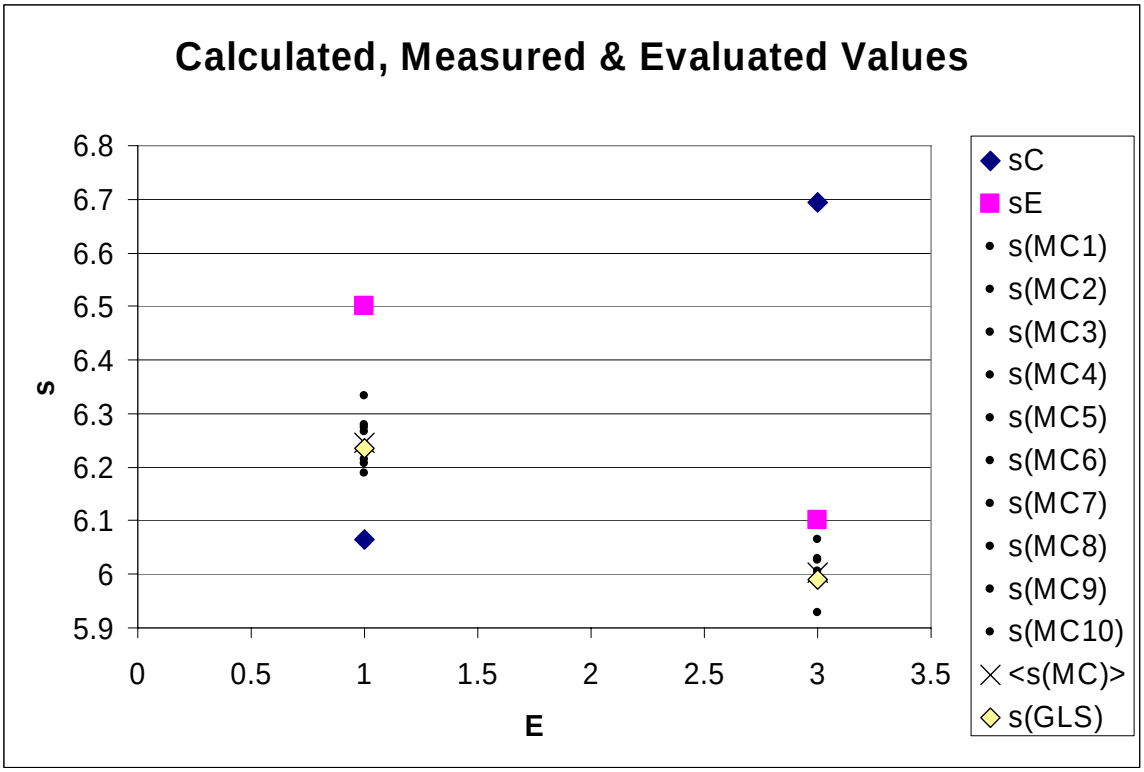


Figure 11: Plot of input data and results obtained from GLS as well as the UMC calculations performed using the "Wide Limits" condition

The results for individual UMC simulations based on only 1000 histories scatter noticeably with respect to the GLS solution. Nevertheless, the tendency toward agreement with the GLS results is clearly evident. When the results from all 10 MC simulations are averaged, the resulting average values are very close to the GLS solution results. **The outcome from this example strongly suggests that the UMC approach described in this report is capable of yielding results that agree very well with GLS under the stated conditions**. The only apparent limitation is the number of sampling histories. The plot in Fig. 11 gives us only a qualitative view of the outcome. The numerical results are summarized in Table 4.

Table 4: Results obtained for Example 4 from GLS as well as the average of 10 UMC calculations with 1000 histories each performed using the "Wide Limits" condition

Quantity	<UMC>	GLS	Difference
$\langle s_1 \rangle$	6.245356668	6.23653	0.14%
$\langle s_2 \rangle$	6.002441299	5.99151	0.18%
$Var(s_1)$	0.072517244	0.075551	-4.18%
$SD(s_1)$	0.267397945	0.274865	-2.79%
% $Err(s_1)$	4.28%	4.41%	-2.96%
$Var(s_2)$	0.076383333	0.0812184	-6.33%
$SD(s_2)$	0.275075323	0.284988	-3.60%
% $Err(s_2)$	4.58%	4.76%	-3.76%
$Cov(s_1,s_2)$	0.058017645	0.0618054	-6.53%
$Cor(s_1,s_2)$	0.772758502	0.789003	-2.10%

Note: The numerical values given here are taken directly from the EXCEL spreadsheet

Other than the fact that the solution value $\langle s_2 \rangle$ is smaller than either of the input values s_{E2} or s_{C2} (perhaps not unusual considering the discrepant input information and strong correlations involved in this exercise), the results from the GLS and UMC approaches appear to be quite reasonable and consistent with each other.

There is one more matter to consider. What happens if the sampling limits are tightened significantly? To explore this possibility, we suppose that the sampling limits on s_1 are (5.4,7.1) while those for s_2 are (5.1,6.9). This is referred to as the "Tight Limits" condition. The 10 sets of UMC calculations described above were repeated with the outcome shown in Fig. 12 and Table 5.

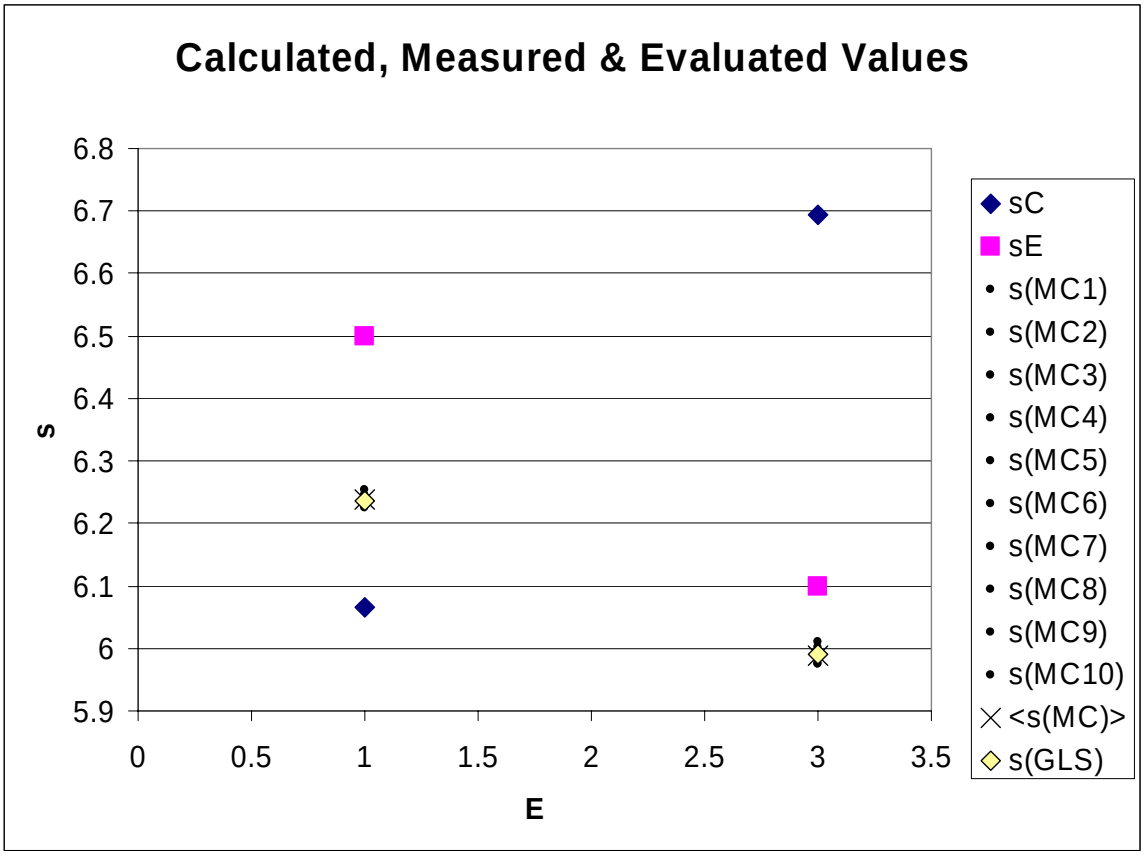


Figure 12: Plot of input data and results obtained from GLS as well as UMC calculations performed using the "Tight Limits" condition

Table 5: Results obtained for Example 4 from GLS as well as the average of 10 UMC calculations with 1000 histories each performed using the "Tight Limits" condition

Quantity	<UMC>	GLS	Difference
$<s_1>$	6.238185665	6.23653	0.03%
$<s_2>$	5.988849267	5.99151	-0.04%
$Var(s_1)$	0.072690798	0.075551	-3.93%
$SD(s_1)$	0.269492943	0.274865	-1.99%
% $Err(s_1)$	4.32%	4.41%	-2.02%
$Var(s_2)$	0.081322156	0.0812184	0.13%
$SD(s_2)$	0.284996909	0.284988	0.00%
% $Err(s_2)$	4.76%	4.76%	0.05%
$Cov(s_1,s_2)$	0.060507895	0.0618054	-2.14%
$Cor(s_1,s_2)$	0.786451732	0.789003	-0.32%

Note: The numerical values given here are taken directly from the EXCEL spreadsheet

The individual results of the 10 separate UMC calculations with 1000 histories all cluster much more closely around the GLS values for the "Tight Limits" sampling condition than they do for the "Wide Limits" sampling condition. Furthermore, the averages of these UMC results are almost identical to the GLS result. This supports our original contention that to achieve good results with the UMC approach it is necessary to find the right combination of sampling limits and MC sampling histories. Skill in doing this can be acquired from experience. The technique of examining convergence of the various computed averages, as considered in this example as well as the preceding ones, is also helpful in obtaining reliable results.

Finally, we apply the test for consistency of the input data as it is described above and reflected in Eq. (17). The outcome for the input information of this example is $\mathcal{X}^2$/(d.o.f.) = 1.7372 (to five significant figures). Therefore, in accordance with the discussion above, we may be justified in multiplying the solution uncertainties by the square root of this number in order to compensate for the discrepancy indicated by this test. This would lead to an increase in the uncertainties derived by the UMC evaluation method by ≈ 30 % to values on the order of 6 %.

5. Conclusions

The numerical examples presented in this report suggest that the UMC method that is described here is viable and can yield results that are consistent with the GLS approach when the conditions that support agreement are satisfied. There was no *a priori* reason to expect that this would not be the outcome since the UMC methodology is quite rigorous. On the other hand, the GLS method is NOT completely rigorous. It relies on approximating the mean values by those which yield the maximum value for the joint probability density function (Maximum Likelihood Assumption). There is also an implicit assumption of linearity throughout. If the probability density function is truly a multivariate normal distribution with respect to the variables being evaluated, and the uncertainties are modest, then the GLS approximation should be quite good. These conditions are frequently satisfied when there is a one-to-one relationship between the experimental and prior data (model-calculated for most evaluations). These conditions are obviously satisfied in Example 4, so the good agreement between the GLS and UMC methods, when the number of traced MC histories is sufficient, had to be anticipated. However, if the "experimental" data included in the evaluation procedure were to involve more complex relationships between the evaluated variables and what is actually measured, *e.g.*, ratios, integral quantities, *etc.*, and possibly large uncertainties and discrepancies are present, then the probability density function could deviate significantly from a normal one and become decidedly skewed. The GLS and UMC solutions would then be expected to differ significantly. Under such circumstances, the UMC solution would be the more defensible one and the GLS solution more suspect.

So, we end this report with the following general conclusions:

(1) The UMC method, while more demanding of computational power than the GLS method, nevertheless appears to be viable. (Actually, much of the computational time in a UMC evaluation will be spent in multiplying matrices.) So, this method should certainly be tested in an actual evaluation exercise. For example, it could be applied to the ^{89}Y evaluation test case that is currently being considered by WPEC Subgroup 24. Investigations such as this would lead to a better understanding of the tradeoffs between large variable sampling volumes and large numbers of MC histories that might need to be traced to achieve adequate convergence of the evaluated quantities (mean values as well as covariance matrices).

(2) Considering the computational capabilities available today to evaluators, it is probably prudent to apply the GLS method of evaluation whenever the circumstances are such that its approximations are sufficiently justified. This approach is much less computationally intensive than the UMC approach and gives precisely reproducible (deterministic) solutions. The conditions suitable for application of the GLS method exist when non-linear effects are modest, uncertainties are not too large, and the included experimental data are directly comparable to the variables being evaluated, in the manner described earlier. However, when it is expected that the probability density function upon which a rigorous evaluation should be based is likely to be noticeably skewed by various factors discussed above, the UMC approach should definitely be applied. Judicious

choices of the sampling limits and numbers of MC sampling histories can potentially optimize this process and insure that it is both practical and worthwhile in spite of the additional computation time required to perform a UMC evaluation.

(3) Computers continue to evolve at a rapid pace, becoming faster and providing more storage capacity at an ever decreasing cost year after year. Eventually, within a very few years, it should become practical to perform ALL fast-neutron reaction cross section evaluations by the UMC method. This approach would insure that biases associated with approximations such as those inherent to the GLS method or various older or even relatively contemporary evaluation methodologies are avoided and rigor is preserved.

Acknowledgements: Without the encouragement, wise suggestions, and considerable effort contributed by Michal Herman (Chairman of WPEC Subgroup 24), Roberto Capote Noy, Andrej Trkov, Soo-youl Oh, Arjan Koning, and Dick McKnight this work would have been impossible. Their stimulating e-mail messages received during the winter months of 2006/2007 were crucial to the development of the UMC nuclear data evaluation concept. Thanks are due also to Nancy Larson who reviewed the sketchy and very difficult to read hand-written notes that formed the nucleus of this report and offered her "blessing" regarding this author's interpretation of some fundamental concepts from probability theory upon which the UMC method is based. The author is indebted to Argonne National Laboratory and to the U.S. Department of Energy for the support it has provided to enable this work to be carried out.

References

[1] Donald L. Smith, *Probability, Statistics, and Data Uncertainties in Nuclear Science and Technology*, American Nuclear Society, La Grange Park, Illinois (1991).

[2] Donald L. Smith, "A Least-Squares Computational 'Tool Kit' ", Report ANL/NDM-128, Argonne National Laboratory (1993).
[http://www.ne.anl.gov/capabilities/nd/reports/ANLNDM121.html]

[3] Satoshi Chiba and Donald L. Smith, "A Suggested Procedure for Resolving an Anomaly in Least-Squares Analysis Known as 'Peelle's Pertinent Puzzle' and the General Implications for Nuclear Data Evaluation", Report ANL/NDM-121, Argonne National Laboratory (1991).
[http://www.ne.anl.gov/capabilities/nd/reports/ANLNDM121.html]

[4] Donald L. Smith, "Covariance Matrices for Nuclear Cross Sections Derived from Nuclear Model Calculations", Report ANL/NDM-159, Argonne National Laboratory (2004).
[http://www.ne.anl.gov/capabilities/nd/reports/ANLNDM141.html]

[5] Donald L. Smith, "Nuclear Data Uncertainties in 2004: A Perspective", *International Conference on Nuclear Data for Science and Technology*, Santa Fe, New Mexico, U.S.A., 26 September – 1 October 2004. Eds. R.C. Haight, M.B. Chadwick, T.

Kawano, and P. Talou, AIP Conference Proceedings No. 769, American Institute of Physics, Melville, New York (2005). See p. 320.

[6] 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).