VARIED APPLICATIONS OF A NEW MAXIMUM-LIKELIHOOD CODE WITH COMPLETE COVARIANCE CAPABILITY

F. Schmittroth

290 6800

Hanford Engineering Development Laboratory
P. O. Box 1970
Richland, Washington 99352

Paper to be presented at seminar workshop, "Theory and Application of Sensitivity and Uncertainty Analysis", to be held in Oak Ridge, Tennessee, August 22-24, 1978.

Hanford Engineering Development Laboratory, Richland, Washington, operated by Westinghouse Hanford Company, a subsidiary of Westinghouse Electric Corporation, for the Department of Energy under contract number EY-76-C-14-2170.

DISCLAIMER

VARIED APPLICATIONS OF A NEW MAXIMUM-LIKELIHOOD CODE WITH COMPLETE COVARIANCE CAPABILITY

F. Schmittroth
Hanford Engineering Development Laboratory
Richland, Washington, USA

ABSTRACT

Applications of a new data-adjustment code are given. The method is based on a maximum-likelihood extension of generalized least-squares methods that allow complete covariance descriptions for the input data and the final adjusted data evaluations. The maximum-likelihood approach is used with a generalized log-normal distribution that provides a way to treat problems with large uncertainties and that circumvents the problem of negative values that can occur for physically positive quantities.

The computer code, FERRET, is written to enable the user to apply it to a large variety of problems by modifying only the input subroutine. The following applications are discussed: A 75-group a priori damage function is adjusted by as much as a factor of two using 14 integral measurements in different reactor spectra. Reactor spectra and dosimeter cross sections are simultaneously adjusted based on both integral measurements and experimental proton-recoil spectra. The simultaneous use of measured reaction rates, measured worths, microscopic measurements, and theoretical models are used to evaluate dosimeter and fission-product cross sections. Applications in the data reduction of neutron cross section measurements and in the evaluation of reactor after heat are also considered.

INTRODUCTION

In recent years, the methods of generalized least-squares[1-6] have been successfully applied to a variety of data-adjustment and unfolding problems; not only in the fields of reactor physics and nuclear data, but in fields as diverse as aerospace and communications. In this study, we discuss a useful extension of these ideas and present several examples that illustrate the method. Brief consideration is also given to the computer-code implementation of the method.

The term, generalized least-squares, connotes two important extensions of the usual least-squares in the present context: the use of complete covariance matrices rather than diagonal matrices to describe both input and output information, and the use of a priori information to obtain solutions for what are otherwise underdetermined problems.

METHOD

Review of Least-Squares

A detailed description of the method will be given elsewhere. Here, we present important features that distinguish it from earlier work and enough detail to discuss subsequent examples.

The usual least-squares equations may be obtained by minimizing[1]

$$S_g = [g_m - g(w)]^t \, M_g^{-1} \, [g_m - g(w)] \tag{1}$$

where g_m is a vector of measured values, and where w is a vector of least-squares parameters that are to be determined. The vector function g(w) may be linearized as $g(w) \cong g(w^*) + \nabla_w g \cdot (w - w^*)$ if g is not already linear. The covariance matrix M_g describes the known uncertainties and correlations in the measured data g_m:

$$(M_g)_{ij} = \mathrm{Cov}\,(g_{mi}, g_{mj}) \; . \tag{2}$$

A priori knowledge of the parameters w can be included by minimizing[2-6]

$$S = S_g + S_o \tag{3}$$

rather than S_g alone where

$$S_o = (w_o - w)^t \, M_{wo}^{-1} \, (w_o - w) \; . \tag{4}$$

The a priori values are denoted by the vector w_o. The uncertainties and correlations in the values w_o are also considered as a priori knowledge and are described by the covariance matrix M_{wo}. With the definitions,

$$f = g_m - g^* \tag{5a}$$

$$x = w - w^* \tag{5b}$$

$$x_o = w_o - w^* \tag{5c}$$

$$M_{xo} = M_{wo} \tag{5d}$$

$$M_f = M_g \tag{5e}$$

the solution to the minimization problem is given by[6]

$$\hat{x}-x_o = M_{xo}\, A^t\, D^{-1}\, (f-Ax_o) \tag{6}$$

where

$$D = A\, M_{xo}\, A^t + M_f \tag{7}$$

and where $\hat{x}$ denotes the "best" estimate of the parameters $x = w-w^*$. The uncertainties and correlations in this estimate are described by the covariance matrix M_x which can be calculated from

$$M_x - M_{xo} = M_{xo}\, A^t\, D^{-1}\, A\, M_{xo} \,. \tag{8}$$

As is well known, the least-squares method is equivalent to a maximum-likelihood approach that makes the assumption of normal distributions.[1] It is also well known that the least-squares equations can be derived without the assumption of normality from the principle of a "minimum-variance unbiased estimator".[1]

Description of A Priori Information and its Inclusion in a Maximum-Likelihood Approach

The description and use of a priori information plays an important role in the present method. As mentioned earlier, a priori knowledge includes covariance information, M_{wo}, as well as a priori values, w_o. A common reason for using a priori data is to provide a unique solution to the least-squares equations in a situation that would be mathematically underdetermined otherwise. However, mathematically overdetermined problems do not negate the practical use of a priori knowledge.

A particular type of a priori information led us to consider a maximum-likelihood approach in contrast to a least-squares approach. Many physical values, especially in reactor physics and nuclear data, are a priori positive. In fact, their physical character frequently dictates that they be described (e.g. plotted) logarithmically. For instance, one might be able to estimate a particular nuclear cross section to within a factor of 2. The a priori value might be 10, for example, with upper and lower bounds of 2x10=20 and 10/2=5, respectively. The factor of 2 uncertainty might itself be quite uncertain; and yet, one might be very certain that the value sought was not a factor of 100 smaller. The least-squares equations do not allow the accurate description of this type of a priori knowledge. The problem is evident in Eq. (4) if one considers a diagonal covariance matrix M_{wo}. The value S_o is a sum of squared residuals and gives a measure of the deviation of w from the a priori values w_o. Clearly, S_o is a symmetric function of w about the values w_o and does not adequately represent the asymmetric limits discussed above. In fact, for large uncertainties as described by M_{wo}, the value of S_o may not even strongly dictate against negative values for w, even though it may be a priori known with complete assurance that w cannot be negative.

An obvious solution is to replace S_o by a term that explicitly accounts for a logarithmic description of the a priori information. Very simple treatments have been considered previously. Specifically, we replace S_o by

$$S_o' = (z_o - z)^t \, M_{zo}^{-1} \, (z_o - z) \ , \tag{9}$$

where

$$z = \ln(w) \tag{10a}$$

$$z_o = \ln(w_o) \ , \tag{10b}$$

so that the minimization of S in Eq. (4) requires that $\ln(w)$ be close to $\ln(w_o)$ in lieu of the requirement that w be close to w_o. Negative values for w are completely ruled out. The covariance matrix M_{zo} in Eq. (9) is related to M_{wo} by

$$\frac{(M_{wo})_{ij}}{\langle w_i \rangle \langle w_j \rangle} = e^{(M_{zo})_{ij}} - 1 \ . \tag{11}$$

Analytic solutions that correspond to Eqs. (6), (7), and (8) are no longer possible. However, Eq. (6) can be used in an iterative fashion to find the desired minimum of

$$S' = S_g + S_o' \ .$$

In line with a logarithmic picture, it is desirable to express uncertainties (covariances) in fractional form as alluded to in Eq. (11). At the same time, we take a step toward a more practical description of the needed covariance matrices. One can show that M_{wo} can be rewritten as

$$\frac{(M_{wo})_{ij}}{\langle w_i \rangle \langle w_j \rangle} = (1+c^2)(1+r_i r_j \rho_{ij}) - 1 \tag{12}$$

where c is a fractional normalization uncertainty (c=1 implies a 100% uncertainty completely correlated for all values $\{w_i\}$). The values $\{r_i\}$ denote any additional point by point fractional uncertainty, and ρ_{ij} is a matrix that describes any further correlations. From Eqs. (11) and (12), we now find

$$M_{zo} = \ln(1+c^2) + \ln(1+r\rho r) \ . \tag{13}$$

Even for relatively large uncertainties, one can usually use

$$M_{zo} \approx c^2 + r\rho r \ . \tag{14}$$

In any event, M_{zo} can be completely specified in terms of fractional as opposed to absolute uncertainties.

A completely rigorous replacement for Eq. (8) is not available. A solution that one can show works well in many cases is given by Eq. (8) with M_{xo} replaced by M_{zo} and with A appropriately redefined to reflect the change from absolute to relative values. For small uncertainties, the method is equivalent to the usual least-squares treatment.

APPLICATIONS

The FERRET Code

The formalism sketched above was implemented in the FERRET computer code.[7] Additional features include maximum use of partitioning to eliminate unneeded subspaces in covariance matrices and to allow sequential use of independent data subsets. Also, indexing techniques were used that make maximum use of available core memory. A secondary advantage of the indexing method is a computational module that can handle vastly different types of problems by simply modifying the input subroutine. Practical experience has demonstrated the algorithms to be reasonably fast and efficient. The examples discussed below were all easily treated with less than 38K with running times on the order of a minute or less on a CDC 6600 computer.

Before proceeding to specific examples, we want to discuss a further parameterization used in the description of the a priori covariance matrix. The computational module of the code does not impose any restrictions on the allowable covariance matrices. However, in practice we have relied on an intuitive parameterization of the covariance matrices for the measured and a priori data. Equations (12-14) already display one step toward this parameterization. As discussed above, the covariance matrices are separated into two components: an overall fractional normalization uncertainty, c, and a second term, $r_i r_j \rho_{ij}$, that describes any additional uncertainties and correlations. In practice, the correlation matrix ρ_{ij}, has been used primarily to describe short-range correlations, and we have further parameterized it by a form such as[6]

$$\rho_{ij} = (1-\theta)\delta_{ij} + \theta e^{\frac{-(i-j)^2}{2\gamma^2}} \qquad (15)$$

where θ denotes the strength of short-range correlations and γ denotes their range. For example, completely uncorrelated data or a priori values are described by $\theta=0$ so that $\rho_{ij}=\delta_{ij}$. The values $\{r_i\}$ are the point-by-point fractional uncertainties.

It is important to emphasize the role of the correlation matrix ρ_{ij} for the a priori covariance matrices. A priori knowledge consists of more than preliminary estimates w_o of least-squares parameters. For

highly underdetermined problems, stiffness or smoothness constraints can be crucial in obtaining final solutions with uncertainties small enough to be of practical value. Any unfolding code must apply some constraint of this nature to obtain useful solutions for this type of problem (e.g., minimization of second derivative terms). Note that ρ_{ij} does not necessarily dictate smoothness, only a particular degree of correlation among neighboring points. An advantage of using covariance matrices to obtain unique solutions is that final uncertainties are obtained that describe the uniqueness of these solutions. If insufficient data or a priori information are supplied, the problem will be apparent from large final uncertainties. The user must then seek additional physical information such as new data values or stronger a priori correlations.

75-Group Damage Function Analysis

The first example is a 75-group damage function unfolding problem chosen for its large uncertainties and adjustments. The damage D_i in a neutron flux ϕ_i is assumed to be given by

$$D_i = \sum_{g=1}^{75} \phi_{ig}\, \sigma_{Dg} \tag{16}$$

where $\{\sigma_{Dg}\}$ are the multigroup values of the damage function $\sigma_D(E)$ being sought. This equation is linearized by the expansion

$$D_i \approx D_i^{(o)} + \sum_g \phi_{ig}^{(o)} (\sigma_{Dg}-\sigma_{Dg}^{(o)}) + \sum_g \sigma_{Dg}^{(o)} (\phi_{ig}-\phi_{ig}^{(o)}) \quad . \tag{17}$$

The flux spectra ϕ_i are assumed to have small uncertainties and are adjusted simultaneously with the damage function $\sigma_D(E)$.

This example included 14 "measured" damage values corresponding to 14 separate spectra and which were computed from a so-called "exact" damage function. These 14 values were then used to adjust an a priori damage function which can then be compared with the "exact" solution. Figure 1 shows the results for the damage function $\sigma_D(E)$.

A principle feature to be noted is the outstanding agreement between the adjusted function and the exact function for neutron energies between about 30 keV and 1 MeV where the fast spectra have a strong response. Thermal spectra were included among the 14 fluxes but Figure 1 centers on the higher energies. The agreement is particularly good in light of adjustments of nearly a decade. In regions of lower sensitivity as near the edges of Figure 1, there is little adjustment of the a priori values, a desirable feature.

Very briefly, a few of the covariance parameters used are as follows. The 14 damage values were assumed to have an uncertainty of 15% independent of each other. The uncertainties on the multigroup fluxes were typically 10-15% for the better known energy groups. No

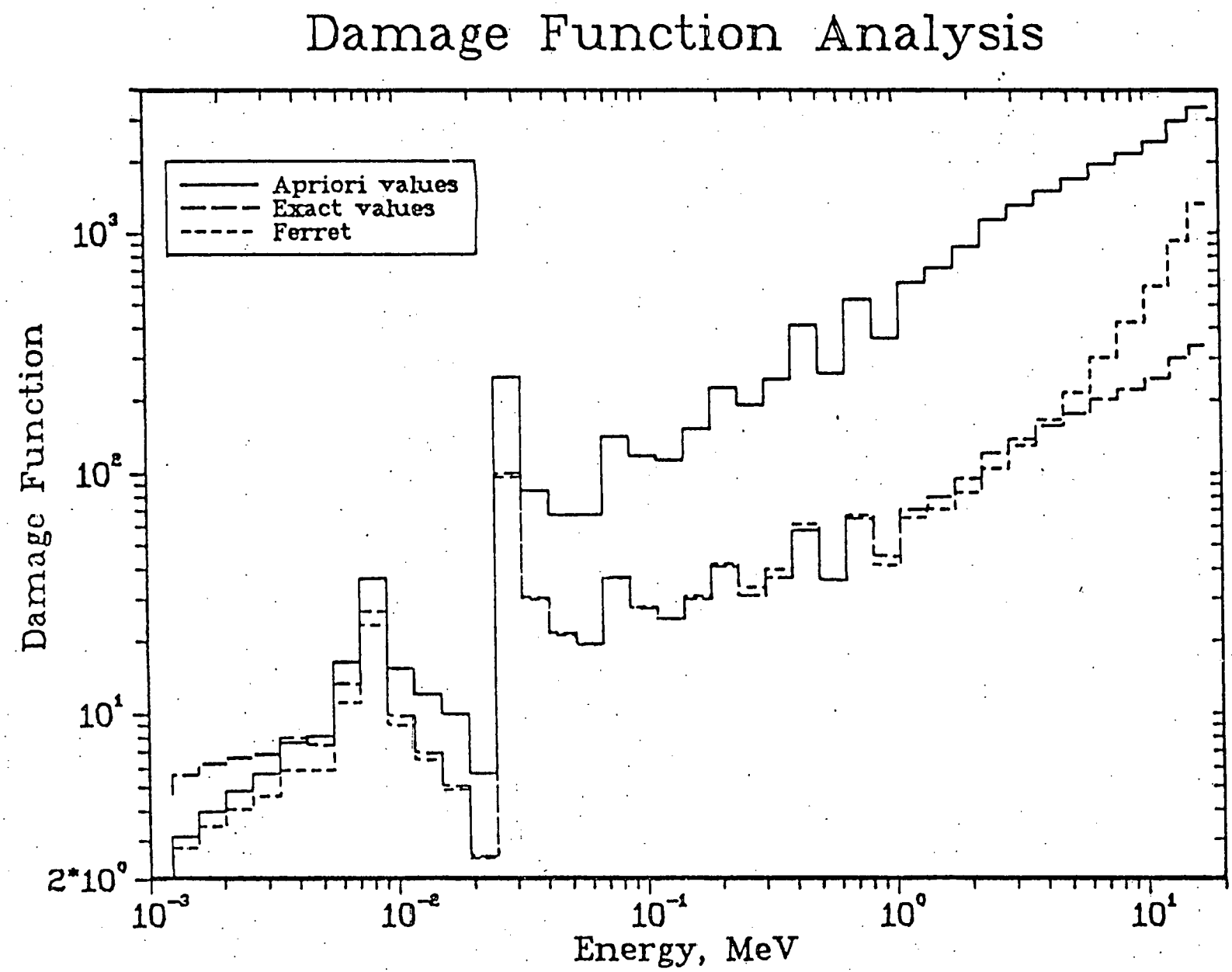


Figure 1. 75-group damage function analysis. Although the analysis extended to thermal energies, only the top 39 groups are shown for clarity.

normalization uncertainty was assumed, and neighboring groups were taken to be partially correlated over about ±2 groups ($c=0$, $r_g \approx .10$-$.15$, $\theta=0.5$, $\gamma=2$). A priori uncertainties for the damage function were taken to be much larger, 1000% normalization uncertainty and 300% group-by-group uncertainties ($c=10$, $r_g=3$); although somewhat stronger short-range correlations were assumed ($\theta=0.9$, $\gamma=5$). One effect of the short-range correlations is readily discernible at the high-energy end of Figure 1 where the adjusted damage function departs in a gradual way from the exact value.

A number of features of the final uncertainties (not illustrated) should be mentioned. Typical uncertainties calculated for the individual groups of the final adjusted damage function are on the order of 300% in those regions of low sensitivity to the integral damage values. This value simply reflects the original 300% a priori group-by-group uncertainty. Of course, even a single integral damage value would suffice to eliminate the 1000% a priori normalization uncertainty. The energy ranges of high sensitivity to integral values are sharply delineated by about an additional factor-of-3 reduction of the final group damage-function uncertainties to about 100%, a value still large in the face of the striking agreement shown in Figure 1. In part, the good agreement in Figure 1 is an artifact of an artificial problem (e.g. in reality, exact damage values cannot be known) and illustrates the potential danger in using such comparisons as a means for testing the accuracy of the final solution. However, although the present algorithm gives reliable uncertainties for more constrained systems or where the uncertainties are smaller, there are indications it overestimates the final uncertainties in loose systems such as the present example. Even so, one can show by direct calculation that all the correlations in the final damage function induced by the integral values are present to the following extent. If one computes the uncertainty in a particular damage value from the adjusted covariance matrix, uncertainties as small as the original "measured" uncertainties are obtained.

Fast-Neutron Spectrum Analysis

A second example we consider is the unfolding or adjustment of fast-neutron spectra based on dosimeter and proton recoil measurements. This example is similar in form to the previous example. It does show, however, how many-channel data (in this case proton-recoil data) can be easily included in the few-channel problem. Dosimeter reaction rates are related to multigroup fluxes and cross sections by

$$R_i = \sum_g \phi_g \sigma_{ig} \tag{18}$$

where i specifies the various dosimeter foils. As an aside, we note that the present code can include multiple spectra (e.g. from different locations within a reactor) at the same time as multiple foils.

This example is based on 12 actual dosimeter measurements in the Engineering Mockup Critical (EMC) as well as proton recoil measurements.[8] It was reported on earlier[9] but without the proton-recoil measurements which are included directly in the analysis here. The 12 dosimeters include resonance detectors, threshold detectors, and broad range detectors. Dosimeter cross sections are based on ENDF/B-IV. The reported reaction rate measurements were typically 3-6% uncertain. The calculation used to generate an a priori flux was assumed to have a normalization uncertainty of $100c=50\%$ and additional group-by-group uncertainties of $100r_g=30\%$ with short-range correlations specified by $\theta=0.5$ and $\gamma=5$. Cross section uncertainties were evaluated and typically were much smaller. The analysis was carried out in 42 energy groups and proton-recoil measurements were included for the high-energy groups 4-22. Proton-recoil uncertainties were taken to be 5% with an additional 10% overall normalization uncertainty.

Results for the neutron flux are shown in Figure 2. Rather than pursue a detailed discussion, we draw attention to one particular feature of this example. Prior to inclusion of the proton recoil results, the neutron flux along with the multigroup representations of the 12 cross section sets were adjusted based solely on the dosimeter reaction rates. As shown by the dotted line in Figure 2, there was a rather noticeable upward adjustment of the neutron flux near 20 keV. Moreover, this adjustment was not confirmed by comparisons with proton-recoil measurements. Therefore, the proton-recoil measurements were included in the same analysis to provide a more direct constraint on the adjusted flux. As seen from the dashed line in Figure 2, the comparatively small uncertainties for the proton-recoil measurements effectively eliminate the previously noted upward flux adjustment. Instead the nuclear cross sections are adjusted, in particular $^{45}Sc(n,\gamma)$ which has a response in this energy region. Figure 3 shows that the scandium cross section had been adjusted upward a little prior to the inclusion of the proton-recoil data. But with the proton-recoil data, the adjustment is greater, more than 9% in spite of assigned uncertainties of only 6% for the a priori cross section in this energy range. There is a strong implication that the scandium cross section should be reexamined.

Multichannel Unfolding

The FERRET code was used to reduce data on recent D-Li neutron yield studies. The data reduction required the subtraction of a background measurement from a foreground measurement and accounted for detector response broadening. Also, although the problem was not underdetermined as in the previous examples, important a priori information was included: a presumed smoothness was specified by the use of a priori short-range correlations, and conservation of energy dictated that the neutron yield was rigorously zero above a known energy. The problem is summarized by Figure 4 which shows both the experimental neutron-yield measurements and the reduced data as a function of the neutron energy. The fluctuations seen in the measurements are due to counting statistics. Relative channel 30 corresponds

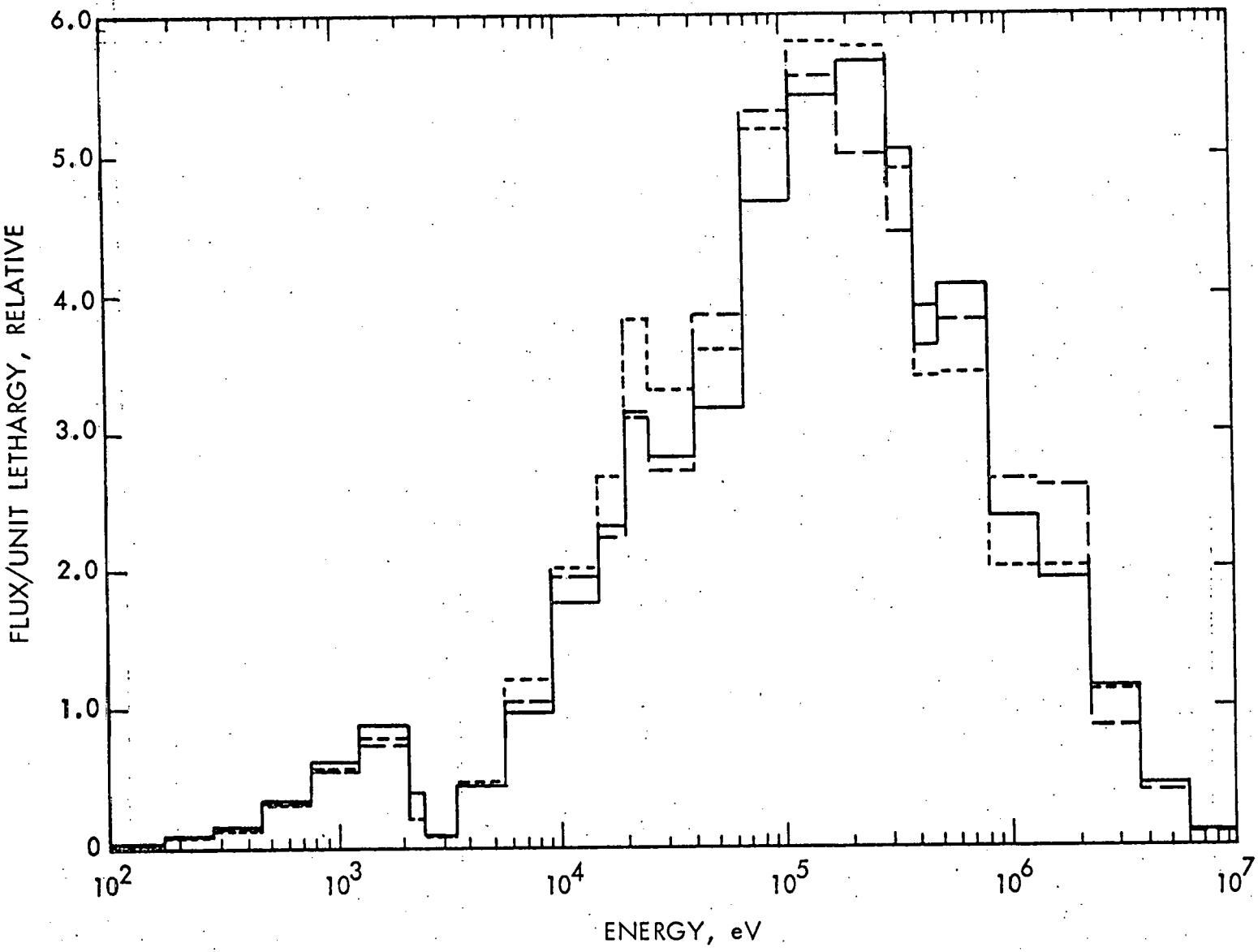


Figure 2. 42-group analysis for the Engineering Mockup Critical (EMC), neutron spectrum: —— a priori value, – – – with proton recoil results, ---- without proton recoil results.

HEDL 7806-13.4

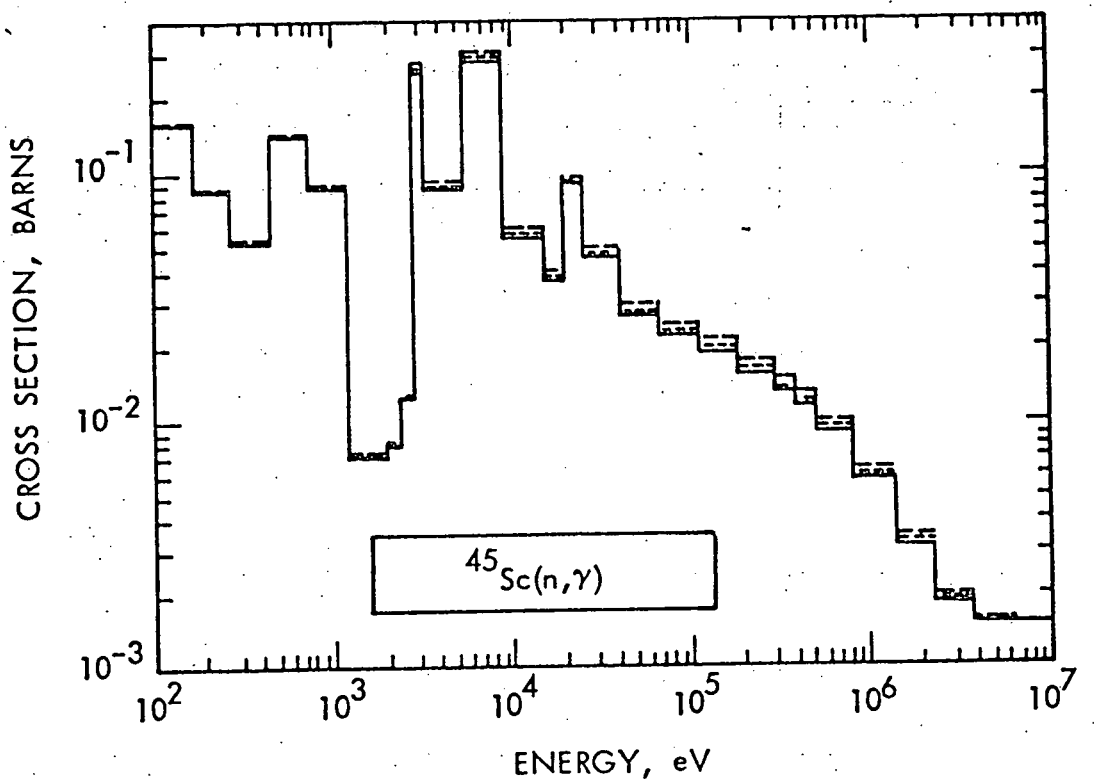


Figure 3. 42-group analysis for the Engineering Mockup Critical (EMC), $^{45}Sc(n,\gamma)$ cross section: —— ENDF/B-IV, — — — with proton recoil results, ---- without proton recoil results.

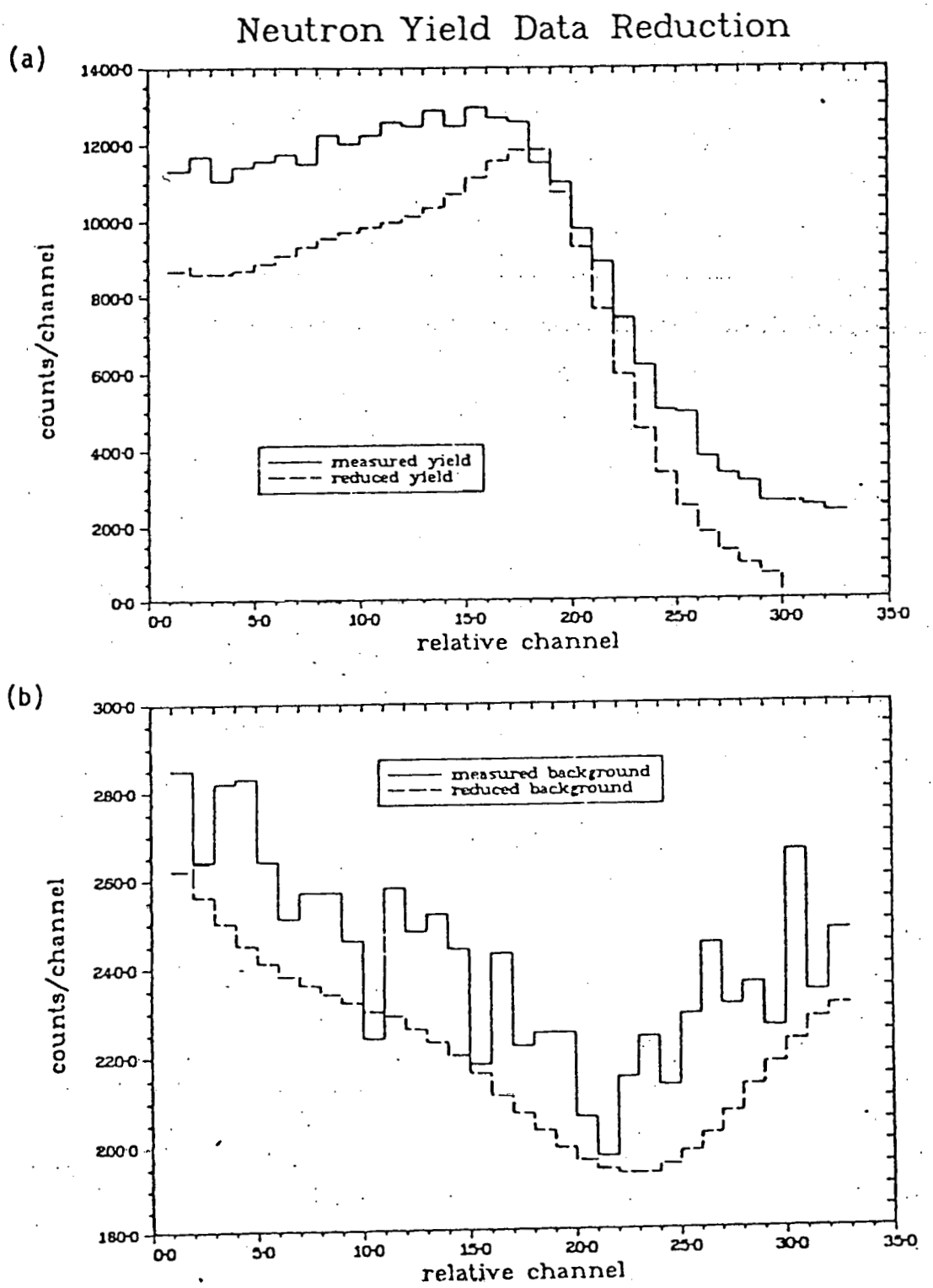


Figure 4. Measured and Evaluated Neutron Yield Data at 12°: (a) Foreground Results; (b) Background Results.

to 50 MeV neutrons, and represents the highest energy for which neutrons are emitted.

A particular advantage of the present method is apparent in Figure 5 which shows a sharp rise in the reduced-yield uncertainties above channel 25. As the foreground values become smaller at the higher energies (channels), one obtains the increased error associated with the background subtraction as the background becomes increasingly important.

Cross Section Evaluation

Figure 6 summarizes the results of an evaluation for the neutron-capture cross section for the fission-product ^{109}Ag. Space precludes a detailed discussion. An important feature of this example is the simultaneous use of different types of integral measurements along with direct microscopic data to obtain a final evaluation. In addition to the microscopic data shown, integral worth measurements from several STEK cores[2,10] and a reaction rate measurement from the Coupled Fast Reactor Measurement Facility[11] (CFRMF) were included. Another feature is the use of a pointwise representation for the cross section in contrast to the multigroup-like representations used in the earlier examples. It is interesting to note that the integral measurements, in a sense, rule in favor of the Weston microscopic measurements. The different microscopic measurements had comparable uncertainties. A final point concerns the character of the uncertainties. By their nature, neutron-capture measurements are often subject to overall normalization errors. This important correlation must be included if a measurement with a relatively large number of data points is not to be unduly weighted. The statistical weighting of the single degree-of-freedom associated with normalization generates less weight than a set of statistically-independent values.

The uncertainties are shown in Figure 6b. For low energies, the reductions in the uncertainties for the adjusted values are due to integral measurements. At the highest energies, the uncertainty reduction to 12% arises from the elimination of a 40% a priori normalization error assumed above 1 keV.

SUMMARY

The examples discussed illustrate the diversity of problems that can be treated by the present method. Their common link is the need to simultaneously consider integral and differential data in a common evaluation. In common with a generalized least-squares approach, final uncertainties and covariances are generated that describe the accuracy of the final results. Particular attention was devoted to the problem of large a priori uncertainties for physical quantities known to be positive.

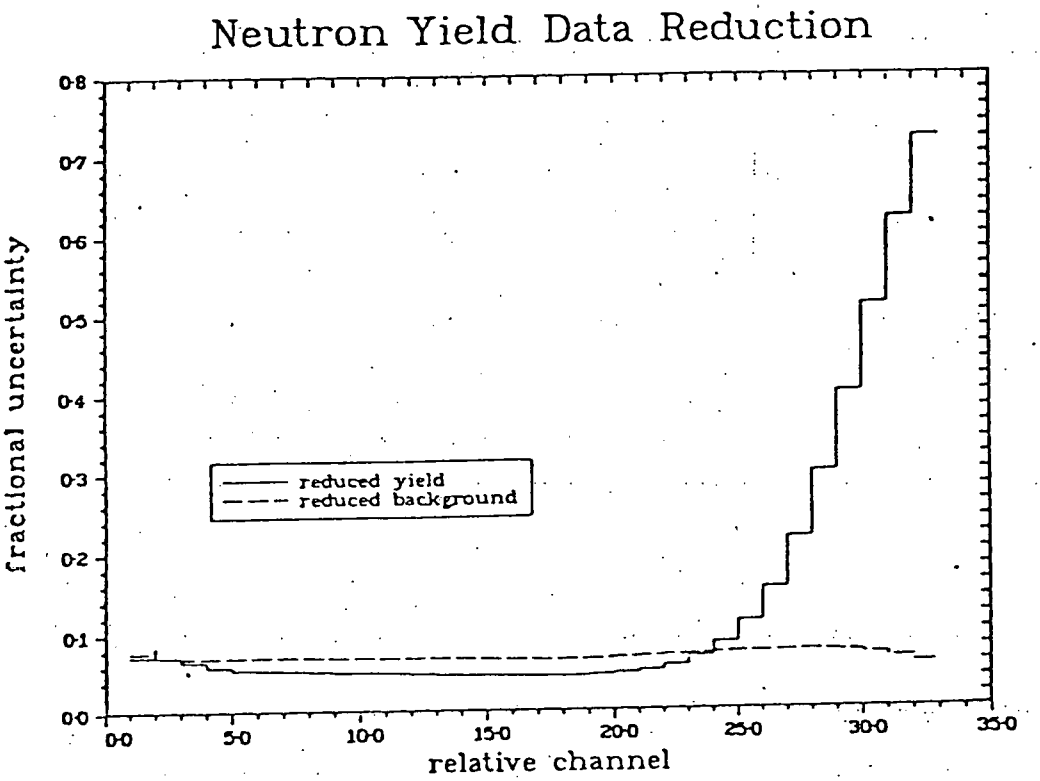
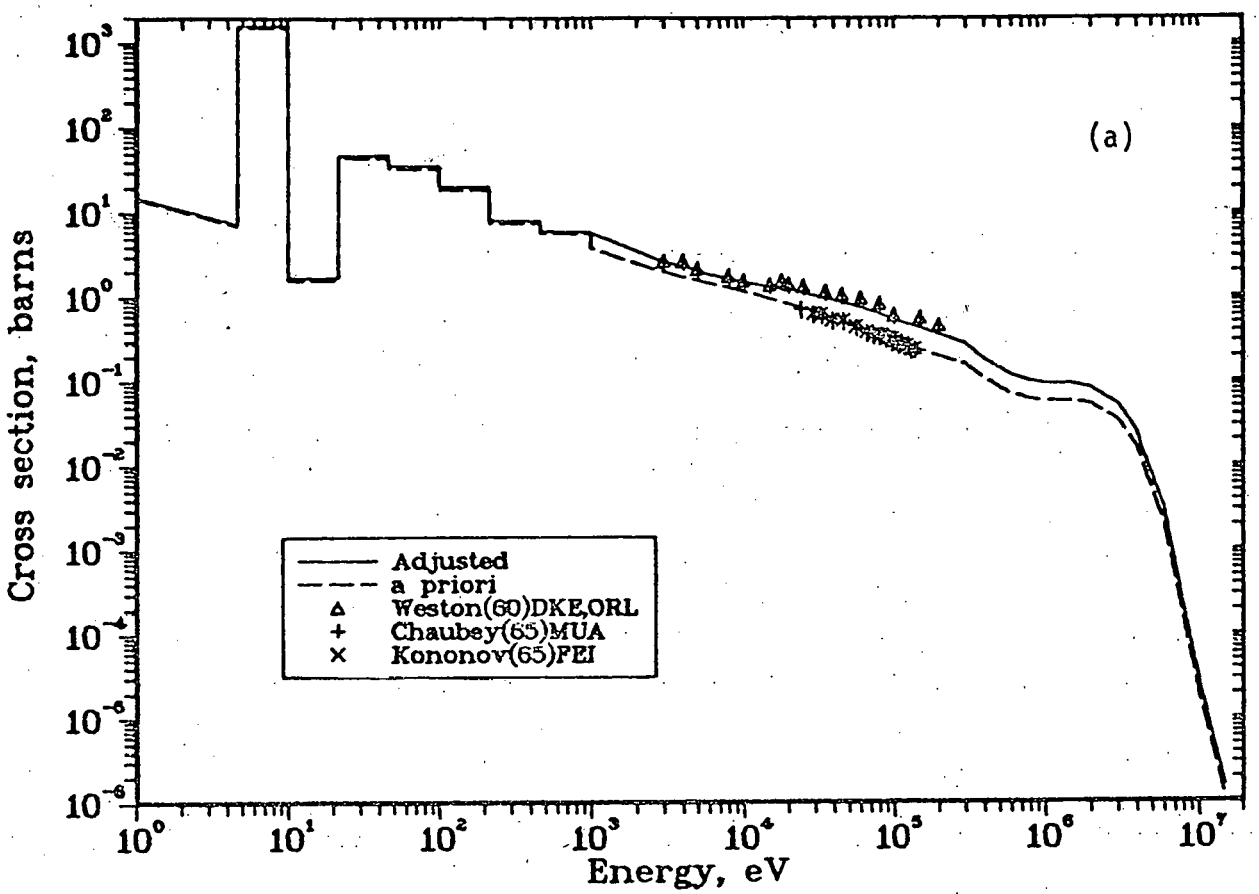


Figure 5. Final uncertainties for reduced neutron-yield data.

Ag 109 evaluation

Case PK 12.2



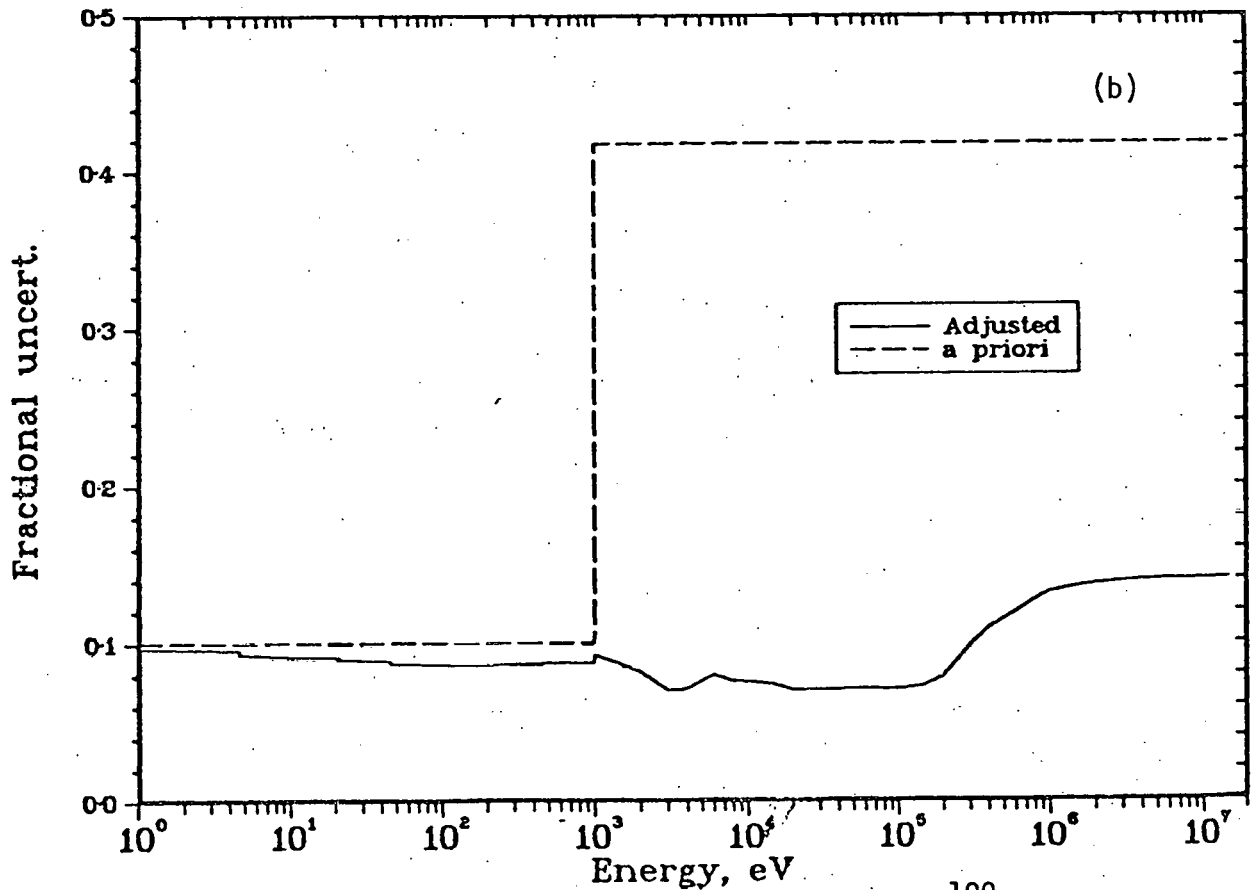


Figure 6, Neutron-Capture cross section evaluation for ^{109}Ag: (a) cross section; (b) uncertainties.

ACKNOWLEDGMENT

The use of preliminary data on recent D-Li neutron yield measurements by D. L. Johnson and F. M. Mann is gratefully acknowledged. The data for the damage function was provided by R. L. Simons. And finally, much of the work on the fission-product example was in conjuncture with R. E. Schenter's group. All the above people are with the Hanford Engineering Development Laboratory (HEDL).

REFERENCES

1. Walter Clark Hamilton, Statistics in Physical Science, The Ronald Press Company, New York (1964).

2. J. B. Dragt, et al., Nucl. Sci. and Eng., 62, 117 (1977).

3. A. Gandini, Nuclear Data and Integral Measurements Correlation of Fast Reactors, Part 2: "Review of Methods", RT/Fi (73) 22, Comitato Nazionale Energia Nucleare (1973).

4. "A Review of Radiation Energy Spectra Unfolding, Proc. of a Seminar-Workshop, April 12-13, 1976", compiled by D. K. Trubey, ORNL/RSIC-40, Oak Ridge National Laboratory, Oak Ridge, TN, October, 1976.

5. F. G. Perey, Least-Squares Dosimetry Unfolding: The Program STAY'SL, ORNL/TM-6062 (ENDF-254), October 1977 and Contributions to Few-Channel Spectrum Unfolding, ORNL/TM-6267 (ENDF-259), February 1978, Oak Ridge National Laboratory, Oak Ridge, TN.

6. F. Schmittroth, Generalized Least-Squares for Data Analysis, HEDL-TME 77-51, Hanford Engineering Development Laboratory, Richland, WA, March 1978.

7. F. Schmittroth, "Methods Development", Core Engineering Technical Progress Report, October, November, December 1977, HEDL-TME 77-49, pp. 51-60, Hanford Engineering Development Laboratory, Richland, WA, April 1978.

8. J. V. Nelson, et al., Measurement and Calculation of Neutron Spectra in the FTR Engineering Mockup Critical, HEDL-TME 76-88, Hanford Engineering Development Laboratory, Richland, WA, Feb. 1977.

9. F. Schmittroth and J. A. Rawlins, "Dosimeter Based Adjustments of Fast-Neutron Spectra", HEDL-SA-1480, Hanford Engineering Development Laboratory, Richland, WA (June 1978); and in ANS Trans., Vol. 28, pg. 621 (June 18-22, 1978).

10. J. W. M. Dekker and H. Ch. Rieff, Adjusted Capture Cross Sections of Fission-Product Nuclides from STEK Reactivity Worths and CFRMF Data, ECN-28, Netherlands Energy Research Foundation (ECN), Petten, The Netherlands (August 1977).

11. Y. D. Harker, et al., Fission Product and Reactor Dosimetry Studies at Coupled Fast Reactivity Measurements Facility, TREE-1259, EG&G Idaho, Inc. (March 1978).