Experimental Uncertainty and Covariance Information in EXFOR Library

N. Otuka¹, R. Capote², S. Kopecky³, A.J.M. Plompen², V.G. Pronyaev³, P. Schillebeeckx³, and D.L. Smith⁴
¹Nuclear Data Section, International Atomic Energy Agency, A-1400 Wien, Austria
²EC-JRC Institute for Reference Materials and Measurements, B-2440 Geel, Belgium
³Institute of Physics and Power Engineering, 249033, Obninsk, Russia
⁴Argonne National Laboratory, 1710 Avenida Del Mundo #1506, Coronado, CA 92118, USA

Abstract. Compilation of experimental uncertainty and covariance information in the EXFOR Library is discussed. Following the presentation of a brief history of information provided in the EXFOR Library, the current EXFOR Formats and their limitations are reviewed. Proposed extensions for neutron-induced reaction cross sections in the fast neutron region and resonance region are also presented.

1 Introduction

Detailed documentation of experimental uncertainty and covariance information is crucial for evaluators who wish to provide their evaluated cross sections with the corresponding covariance matrices based on the experimental knowledge. A typical example is uranium-235 and other actinide neutron-induced fission cross section evaluations performed by least squares fitting to experimental cross sections. Evaluators may use generalized least square fitting codes developed for nuclear data evaluation (e.g., GMA [1], SOK [2]), and can obtain the evaluated cross sections from experimental cross sections and their covariances. Also various approaches used to combine covariances of experimental results and nuclear model calculation results have been proposed [3]. Detailed experimental uncertainty information can often be lengthy and thus not suitable for some types of publications (e.g., journals, conference proceedings). Consequently, the EXFOR Library [4], maintained by the International Network of Nuclear Reaction Data Centres (NRDC) [5] under the auspices of the International Atomic Energy Agency Nuclear Data Section (IAEA-NDS), can serve to provide storage for the relevant information.

In 1978, F.G. Perey showed the importance of documentation [6] by referring to the 2080 keV resonance energy in the neutron induced reaction on carbon measured at Harwell [7]. In this experiment, two energies (2078.31 and 2079.2 keV) were obtained from two time-of-flight paths (50 and 100 m). The independent and common parts of the estimated uncertainties in two flight-paths and two time-of-flight measurements are missing in the original publication. Fortunately, Perey was able to obtain estimates of these partial uncertainties personally from the author, and could construct a 2×2 covariance matrix. The off-diagonal element of the matrix lies between two diagonal elements (i.e., \( V_{11} < V_{12} < V_{22} \)), and the resultant least squares solution (2078.27 keV) is lower than the two measured resonance energies thereby demonstrating the PPP phenomenon (Peelle’s Pertinent Puzzle). Perey concluded as follows: “We urge experimentalists to report the uncertainties in their measurements in such a fashion that the covariance matrix of their results can be generated. We hope that data compilers in the future will expand their data compilation formats such that this valuable information can be made available”.

Now, three decades later, neutron-induced reaction cross sections are often evaluated along with their covariances for general purpose evaluated data libraries. A standardized Guide to the Expression of Uncertainty in Measurement (GUM) [8] has been prepared to provide useful information about estimating uncertainties in data. Also, good text books are available that deal particularly with nuclear data covariances [9-10]. However, the current contents of the EXFOR Library are still not satisfactory for evaluation applications. Recently, at the IAEA Technical Meeting on Neutron Cross Section Covariances (September 2010 [11]), data centres received various recommendations from experts familiar with experiments and evaluations on how to improve the situation, and EXFOR compilers are urged to address this matter in collaboration with experimentalists.

In this report, limitations of the current EXFOR Formats [12] and their extensions for documentation of experimental uncertainties and covariances are discussed for neutron-induced reaction data in the resonance and fast-neutron regions. Detailed discussions with more...
examples will be published separately [13]. Note that a portion of the extended formats introduced in this article is still under consideration and will be discussed further at the next NRDC Meeting (April, 2012).

2 Current EXFOR Formats

In 1980, M. Bhat submitted a working paper entitled “Proposal to include detailed information on systematic errors in the international data exchange files” to the 1980 NRDC meeting [14]. The proposal was adopted after a slight modification in 1982 [15] and it has been continuously used until now by the EXFOR compilers without major revisions.

A coding sample based on the format agreed upon at this meeting is shown in Fig. 1. In this example, two constant “systematic” uncertainties \( \varepsilon_1(E) \) and \( \varepsilon_2(E) \) are coded under the data headings `ERR-1` and `ERR-2`. The incident energy \( E \), and its uncertainty \( \Delta E \), are coded with the corresponding cross section \( \sigma(E) \) under the data headings `EN`, `EN-ERR` and `DATA`, respectively. The energy dependent “statistical” uncertainty \( \varepsilon_s(E) \) and “systematic” uncertainties \( \varepsilon_i(E) \) and \( \varepsilon_3(E) \) are coded under the data headings `ERR-S` and `ERR-3` with the energy dependent total uncertainty \( \varepsilon_T(E) \) under `ERR-T`. For the last component `ERR-4`, only the lower and upper boundaries \( \varepsilon_{4\text{LM}} \) and \( \varepsilon_{4\text{HM}} \) are coded. All partial uncertainties must be defined under the keyword `ERR-ANALYS`.

![Fig. 1. Current EXFOR format for uncertainties.](image)

The major data headings available for specifying uncertainties of cross sections are summarized in Table 1. Note that only uncertainties that are \( 1\sigma \) standard deviations are coded under these headings, and other values (e.g., \( 2\sigma \)) are converted to \( 1\sigma \) values before entering these data into the EXFOR Library.

Experimentalists often report their cross sections with the “total” uncertainties and fail to specify partial uncertainties. These total uncertainties are useless for covariance analysis, and the compiler is asked to keep the uncertainties under the data heading `DATA-ERR` (i.e., uncertainties for which details are unknown). However, if the EXFOR compiler is fortunate enough to find all the partial uncertainties as well as the total uncertainties, then the total uncertainties are coded under the data heading `ERR-T` with their components under `ERR-1`, `ERR-2` etc. In the past, the importance of the partial uncertainties was not well recognized by some EXFOR compilers, and sometimes they kept only the total uncertainties even if the authors gave partial uncertainties in their report.

### Table 1. Major data headings for uncertainties (present).

| Heading | Definitions |
|---------|-------------|
| `ERR-T` | Total uncertainty |
| `ERR-S` | Statistical uncertainty |
| `ERR-SYS` | Total systematic uncertainty |
| `ERR-1` | 1st systematic uncertainty |
| `ERR-2` | 2nd systematic uncertainty |
| `DATA-ERR` | Data uncertainty (details unknown) |
| `MONIT-ERR` | Uncertainty in monitor cross sections |

Sometimes EXFOR compilers find reports which are well-suited for covariance evaluation (e.g., [16-17]). Unfortunately the current EXFOR formats cannot accommodate all the information given in such reports. In general, the neutron cross section covariance between two neutron energies \( E_i \) and \( E_j \) is expressed in terms of the correlation coefficients and partial uncertainties as follows:

\[
\text{cov}(\sigma_i, \sigma_j) = \sum_{k=1}^{n} \text{cov}_{k} (\sigma_i, \sigma_j)
\]

\[
= \sum_{k=1}^{n} c_{kij} \Delta \sigma_i \cdot \Delta \sigma_j
\]

where \( \text{cov}(\sigma_i, \sigma_j) \) and \( \text{cov}_{k} (\sigma_i, \sigma_j) \) are the total covariance and partial covariance due to the \( k \)-th source of uncertainty, \( c_{kij} \) is the micro-correlation coefficient of the \( k \)-th partial uncertainty between the two cross sections \( \sigma_i \) and \( \sigma_j \), and \( \Delta \sigma_i \) stands for the \( k \)-th partial uncertainty for the cross section \( \sigma_i \). This equation reveals that the evaluators need the correlation coefficients \( c_{kij} \) in addition to the partial uncertainties \( \Delta \sigma_i \). For some specific sources of uncertainties, evaluators can assign a proper value to \( c_{kij} \). For instance, uncertainties coded under the data heading `ERR-S` may be always treated as uncorrelated (i.e., \( c_{kij} = 0 \)). Uncertainty due to the intensity of a gamma line adopted from a decay data table for activation cross section determination can always be regarded as fully correlated (i.e., \( c_{kij} = 1 \)). In many cases, however, it is not a trivial task to assign proper values to \( c_{kij} \) without further experimental information. Correlation due to the sample mass may be treated as fully correlated.
only when the evaluator knows the same sample was used at all incident energies. When a correlation coefficient is a constant for a given source, the value can be coded under the keyword ERR-ANALYS in accordance with the current rule. But this information has rarely been given by authors, and only a few EXFOR entries adopt this option for compilation of cross sections (e.g., those measured at the ANL fast neutron generator (FNG) [18-19]).

Experimental total covariances or corresponding macro-correlation coefficients,

\[ c_{ij} = \text{cov}(\sigma_i, \sigma_j)/(\Delta \sigma_i \Delta \sigma_j), \]  

(\Delta \sigma_i \text{ is the total uncertainty of } \sigma_i) can be coded under the keyword COVARIANCE. In the current format rule, however, the matrix elements are treated as free text and therefore are not suitable as inputs to computer programs, especially when the matrix is huge (e.g., covariances of high resolution time-of-flight data in the resolved resonance region). Also, updating of the matrix to account for changes of the partial uncertainties (e.g., revision of the standard cross section with its covariance) is not easy if covariance matrices of partial uncertainties are not explicitly given in the EXFOR Library.

3 Extensions of the EXFOR Formats

In order to improve the capability of the EXFOR Formats to adequately represent cross section data uncertainties and covariances, various extensions of EXFOR formats are under consideration. Hereafter, we limit our discussion to neutron cross sections in the fast neutron and resonance regions. The reader can refer to [13] for more details.

3.1. Covariances in the Fast Neutron Region

As discussed in the previous section, the EXFOR formats must be extended to record the correlation properties of the partial uncertainties. The following two extensions were proposed in the last NRDC Meeting (Vienna, May 2011) [20].

1. An indication should be provided for the correlation property for each source of the partial uncertainty by flags (\( \tau \) for “fully correlated”, \( \Upsilon \) for “uncorrelated”, \( \Pi \) for “partially correlated”, \( \zeta \) for “correlated” where the range is unknown). Note that \( \zeta, \Upsilon \) and \( \Pi \) are also known as “long energy range correlation (LERC)”, “short energy range correlation (SERC)” and “medium energy range correlation (MERC)” in neutron-induced reaction cross section uncertainties [10].

2. The total and partial covariances (or corresponding macro and micro-correlation coefficients) should be included in computer-readable form.

The EXFOR users are able to derive the full covariance matrix if the experimentalists provide all the necessary information, and the EXFOR compilers record it in this extended format.

A coding sample for the fast neutron region is given in Fig. 2. This sample shows a compilation of the \(^{241}\text{Am}(n,2n)^{240}\text{Am}\) cross section measured at the IRMM Van de Graaff laboratory [17,21]. In this example, the total uncertainty and 9 partial uncertainties (MONIT-ERR, ERR-1, ERR-2,..., ERR-8) are defined under the keyword ERR-ANALYS with their correlation properties (\( \tau, \Upsilon, \Pi \)). There are 2 partially correlated uncertainties (MONIT-ERR, ERR-5), and their micro-correlation coefficients are coded under the keyword COVARIANCE explicitly, along with the macro correlation coefficients.

Fig. 2. Coding sample of uncertainties and covariances using the extended EXFOR formats (fast neutron region).

Fig. 3 compares macro-correlations \( c_{ij} \) derived from the full information of micro-correlation (upper panel) provided by the authors with the results obtained by
assuming the total systematic error is fully correlated (lower panel). In other words, in the latter case the assumption is adopted that $c_{ij}=1$ except for the uncertainty due to counting statistics. The main correlations are well represented by the simplifying assumption for this specific example.

There has been considerable variety in the uncertainty terminology used by experimentalists, and this makes it difficult to achieve consistency in EXFOR entries. EXFOR has categorized uncertainty components as “statistical” (ERR-S) and “systematic” (ERR-1, ERR-2,...), but this approach is not recommended in GUM [8]. From the perspective of evaluating uncertainties and covariances, the specification of micro-correlation properties ($c_{ij}$) is crucial, as is evident from Eq. (1), and the current formats must store uncorrelated component under ERR-S and correlated components under ERR-1, etc, effectively in line with [8]. In the extended formats, the correlation properties will be indicated by the new flags (F, U, P, C) while the data headings will function as indicated in Table 2.

### 3.2. Covariance in the Resonance Region

Covariances of time-of-flight spectra $Z$ (transmission, reaction yield, self indication) are fundamental experimental information needed to evaluate covariances of resonance parameters. The number of high resolution time-of-flight bins $m$ is typically $10^4$. Its covariance matrix becomes huge, and it is not realistic to store all partial covariance matrices (or corresponding micro correlation coefficients) in the EXFOR Library. Clearly, a different approach is required.

One approach that can be taken under these circumstances involves what is known as a Cholesky decomposition. In general, the full covariance matrix $V_z$ for a spectrum $Z$ obtained from an analysis involving a parameter vector $\mathbf{\alpha}$ with dimension $n$ is expressed as

$$ V_z = M_z M_z^T + S_{\alpha} \mathbf{\alpha} \mathbf{\alpha}^T, $$

where $M_z$ is a diagonal $m \times m$ matrix consisting of uncorrelated uncertainty components, while $S_{\alpha}$ ($n \times n$ matrix) and $S_{\alpha} = E_z \mathbf{\alpha} \mathbf{\alpha}^T$ ($m \times n$ matrix) are the covariance matrix and the sensitivity (functional) matrix of $Z$ to $\mathbf{\alpha}$. Because the covariance matrix $V_{\alpha}$ is positive definite, there is an $n \times n$ triangular matrix $L_{\alpha}$ which satisfies

$$ V_{\alpha} = L_{\alpha} L_{\alpha}^T $$

(Cholesky decomposition). Finally, the covariance matrix can be written as

$$ V_z = M_z M_z^T + D_{\alpha} D_{\alpha}^T $$

with an $m \times n$ matrix $D_{\alpha} = S_{\alpha} L_{\alpha}$. Such a decomposition of the covariance matrix offers an efficient tool to compile time-of-flight covariance matrices in EXFOR (i.e., the AGS format [22-25]). Actually $M_z$ and $D_{\alpha}$ are $m \times 1$ and $m \times n$ matrices, and they can be treated as typical partial uncertainties of $(1+n)$ sources at $m$ neutron energies in the EXFOR formats.

### Table 2. Major data headings for uncertainties (future).

| Heading     | Definitions                        |
|-------------|------------------------------------|
| ERR-T       | Total uncertainty                  |
| ERR-S       | Total statistical uncertainty      |
| ERR-SYS     | Total systematic uncertainty       |
| ERR-1       | 1st partial uncertainty            |
| ERR-2       | 2nd partial uncertainty            |
| DATA-ERR    | Data uncertainty (details unknown) |
| MONIT-ERR   | Uncertainty in monitor cross sections |

**Table 2. Major data headings for uncertainties (future).**
Fig. 4. Coding sample of uncertainties and covariances in the extended EXFOR formats (resonance region). Only two lines of a spectrum of 25 288 time-of-flight bins are shown.

A coding sample for the covariances of time-of-flight spectra in the resonance region is shown in Fig. 4, where the Cd+n transmission spectra with their uncertainties measured at the IRMM GELINA facility [26] are given. Uncertainties coded under ERR-S and ERR-1 (i = 1 to 4) in this coding sample correspond to \( M_i \) and \( D_{ij} \) in Eq. (4), respectively, and the EXFOR data users can treat them as uncorrelated (\( c_{k,ij} = 0 \)) and fully correlated (\( c_{k,ij} = 1 \)) uncertainties in Eq. (1).

4 Summary

The present paper has discussed the current EXFOR formats, as well as their extension to enable better documentation and more complete representation of experimental uncertainties and covariances than was possible with the existing scheme. In order to provide useful information on experimental uncertainties and covariances to neutron cross section evaluators and other EXFOR data users, within the framework of the extended formats, experimentalists are encouraged to provide detailed information on energy dependent partial uncertainties and their correlation properties. Primary publications of experimental works published as journal articles and conference proceedings are often not suitable to record such detailed information because the numerical data tables may be too lengthy. Reports published by laboratories or data centres (e.g., the INDC report series published by IAEA Nuclear Data Section) may be adequate to publish such supplemental numerical data tables and other information which can be included in the EXFOR entries. Close collaboration between experimentalists and EXFOR compilers may be required for this to be feasible in practice. Furthermore, the development of useful software tools to process information compiled in the extended formats is important. To this end, the IAEA Nuclear Data Section is developing a web-based system to construct and visualize macro-correlation coefficients based on partial uncertainties compiled in the EXFOR Library [27].

Discussions with T. Kawano (LANL) and K. Kobayashi (KURRI) are acknowledged. One of the authors (NO) is indebted to V. Semkova, S. Simakov and R. Forrest (IAEA Nuclear Data Section) for their discussion and encouragement.

References

1. W.P. Poenitz, S.E. Aumeier, Report ANL/NDM-139 (Argonne National Laboratory, 1997)
2. T. Kawano, H. Matsunobu, T. Murata, A. Zukeran, Y. Nakajima, K. Kawai, O. Iwamoto, K. Shibata, T. Nakagawa, T. Ohsawa, M. Baba, T. Yoshida, Report JAERI-Research 2000-004 (Japan Atomic Energy Research Institute, 2000)
3. R. Capote, D.L. Smith, A. Trkov, EPJ Web of Conf. 8 (2010) 04001, and references therein
4. N. Otuka, S. Dunaeva, E. Dupont, O. Schwerer, A. Blokhin, J. Korean Phys. Soc. 59 No. 2, (2011) 1292, and references therein
5. N. Otuka, S. Dunaeva (ed.), Report INDC(NDS)-401 (Rev.5) (International Atomic Energy Agency, 2010)
6. F.G. Perey, Proc. Int. Conf. on Neutron Physics and Nuclear Data for Reactors and Other Applied Purposes (Harwell CONF-780921, 1978) 104
7. G.D. James, Proc. Int. Symp. on Neutron Standards and Applications (NBS Special Publication 493, 1977) 319
8. Joint Committee for Guides in Metrology (JCGM), Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM) (JCGM 100:2008)
9. W. Mannhart, Report INDC(NDS)-0588 (International Atomic Energy Agency, 2011)
10. D.L. Smith, Probability, statistics and data uncertainties in nuclear science and technology (American Nuclear Society, La Grange Park, IL, USA 1991)
11. A. Trkov, D.L. Smith, R. Capote, Report INDC(NDS)-0582 (International Atomic Energy Agency, 2011)
12. N. Otuka (ed.), Report IAEA-NDS-207 Rev.2011/01 (International Atomic Energy Agency, 2011)
13. D.L. Smith, N. Otuka, Nucl. Data Sheets (to be published in 2012)
14. G.J. Wyant, A.M. Daly (ed.), Report INDC(NDS)-125 (International Atomic Energy Agency, 1981)
15. M. Lammer, O. Schwerer (ed.), Report INDC(NDS)-141 (International Atomic Energy Agency, 1982)
16. D.L. Smith, J.W. Meadows, I. Kanno, Report ANL/NDM-85 (Argonne National Laboratory, 1984); EXFOR 12904
17. C. Sage, Mesures des sections efficaces totale et \((n,2n)\) sur \(^{241}\)Am (Université de Strasbourg, 2009)
18. G. Winkler, D.L. Smith, J.W. Meadows, Nucl. Sci. Eng. 76, (1980) 30; EXFOR 10921
19. P. Guenther, D.L. Smith, A.B. Smith, J.F. Whalen, Nucl. Phys. A448, (1986) 280; EXFOR 12869
20. N. Otuka (ed.), Report INDC(NDS)-0593 (International Atomic Energy Agency, 2011)
21. C. Sage, V. Semkova, O. Bouland, P. Dessagne, A. Fernandez, F. Gunsig, C. Naestren, G. Noguere, H. Ottmar, A.J.M. Plompem, P. Romain, G. Rudolf, J. Somers, F. Wastin, Phys. Rev. C 81, (2010) 064604; EXFOR 23114
22. C. Bastian, Proc. Int. Symposium on Nuclear Data Evaluation Methodology, Brookhaven National Laboratory, USA (1992) 642
23. C. Bastian, Proc. Int. Conf. on Neutron Research and Industry, Creta, Greece (1996) 611
24. C. Bastian, A. Borella, F. Gunsing, J. Heyse, S. Kopecky, G. Noguere, P. Siegler, P. Schillebeeckx,
25. N. Otuka, A. Borella, S. Kopecky, C. Lampoudis, P. Schillebeeckx, J. Korean Phys. Soc. 59 No. 2, (2011) 1314

26. S. Kopecky, I. Ivanov, M. Moxon, P. Schillebeeckx, P. Siegler, I. Sirakov, Nucl. Instrum. Meth. B267, (2009) 2345; EXFOR 23077

27. V. Zerkin (these proceedings)