EPJ Web of Conferences **242**, 05001 (2020) https://doi.org/10.1051/epjconf/202024205001
FPY 2019

From fission yield measurements to evaluation: new statistical methodology applied to ^{235}U(n_{th}, f) mass yields

Grégoire Kessedjian[1,*], *Sidi Mohamed* Cheikh[1], *Abdelazize* Chebboubi[2] and *Olivier* Serot[2]

[1]LPSC, Université Grenoble-Alpes, CNRS/IN2P3, F-38026 Grenoble Cedex, France
[2]CEA, DEN, DER, SPRC, LEPh, Cadarache center, F-13108 Saint Paul lez Durance, France

Abstract. The study of fission yields has a major impact on the characterization and understanding of the fission process and is mandatory for reactor applications. The mass and isotopic yields of the fission fragments have a direct influence on the predictions of fuel burn-up and decay heat. Moreover, these data are requested for other studies as delayed neutron evaluation, antineutrino flux assessment or reactor program. Today, the lack of covariance matrix associated to evaluated fission yields induces overestimated uncertainties of mass yields since these observables result from the sum of isotopic and isomeric yields. Our collaboration starts a new program in the field of the evaluation of fission products in addition to the current experimental program. The goal is to define a new methodology of evaluation based on statistical tests in order to provide the best estimation with consistent sets of measurements. A ranking of solutions with associated covariance based on Shannon's entropy criterion is proposed for the mass yields from ^{235}U(n_{th}, f) reaction.

1 Introduction

Fission yields evaluation represents the synthesis of experimental and theoretical knowledges in order to perform the best estimation of mass, isotopic and isomeric yields. Nevertheless the estimation of thes observables are drastically based on experimental data since the modelling of fission process is not predictive. Today, the output of fission yields evaluation is available as a function of isotopic and isomeric yields. As a consequence, mass yields are the sum of isobar nuclei and their quadratic sum to deduce uncertainties. So without any correct covariance matrix, mass yields uncertainties are greater than isotopic yields. This consequence is in contradiction with experimental knowledges where the abundance of mass yields measurements is clearly dominant. Thus, we expect the uncertainties on this latter observable to be lower than those of isotopic yields. Covariance matrix assessment depends on the evaluation process and its validity assumes that all measurements are statistically in agreement. These last years, different covariance matrices have been suggested but the experimental part of those are neglected in covariance evaluation [1][2][3][4] or applications [5][6]. In the first part we present an assessment methodology based on statistical test. The consideration of experimental data is crucial in the definition of the covariance of evaluations where models are not predictive. A large range of data are listed in the EXFOR data bank but a lot of them cover partial mass range. Data are also provided for different incident neutron

*e-mail: kessedjian@lpsc.in2p3.fr

Table 1. Numbers of common masses measured $n_{\hat{A}}$ for the 8 datasets selected.

Datasets	Cumulative measurement number	Cumulative mass number	Index	0	1	2	3	4	5	6	7
Maeck78	37	37	0	37							
Thierens75	74	48	1	26	37						
Diiorio77	138	64	2	37	36	64					
Bail07	168	67	3	17	19	28	30				
Tsoukatos68	232	67	4	37	36	64	28	64			
Tsoukatos68b	260	67	5	23	20	28	13	28	28		
Rosman83	267	72	6	0	0	1	2	1	0	7	
Mathiews83	279	77	7	0	0	1	2	1	0	7	12

energy with different mass resolutions. The merge of data could generate the non-unique solutions according to consistent datasets and will be presented in the second part. In the last part, a ranking of solutions with associated covariance is proposed according to the Shannon's entropy criterion.

2 Experimental datasets

For this first work, the development of the evaluation method is focused on ^{235}U(n_{th}, f) reaction because of a large number available measurements. Thus, for this reaction the EXFOR [7] database permits to cover all the mass range of fission products. Absolute mass yields Y(A) is obtained using the self-normalization of this observable according to the equation:

$$\sum_{A} Y(A) = \Omega \tag{1}$$

where Ω is the normalization factor, here $\Omega = 2$. Unfortunately, for most of these measurements, only statistical uncertainties is provided and systematic uncertainties are estimated in the best cases. Thus the estimation of covariance matrix is non-obvious and represents an important work which will be described in a future work.

2.1 Statistical test on the compatibility of available data

In order to reduce the data and to merge all the measurements for a given mass, it is necessary to test the compatibility of the data. Two kind of data are presented: i) the full range mass yield measurements but not necessarily with sufficient mass resolution ($3\sigma_A < 1amu$), ii) uncomplete mass range generating relative measurements or relative normalizations by the authors. Through the EXFOR [7] database, we chose to test the methodology only on 8 datasets: W.J. Maeck et al. [8], G. Diiorio et al. [9], H. Thierens et al. [10], A. Bail et al. [11], M.P. Tsoukatos [12] with two different datasets, K.J. Rosman [13], C.K. Mathiews [14]. These data correspond to 279 measurements over 77 masses (from A=77 to A=154 with a lack of A=122 in the data used in this work). With this selection, we cover both peaks, allowing the absolute normalization of our evaluation. Thus, assuming independent Gaussian distributions associated to the measurements without explicit information on correlation data, we can calculate the χ^2 using the n_A common measured mass number. This value is compared to the limited χ^2 value (χ^2_{lim}) given for a 99.5% confidence level. In practice, we calculate the

Table 2. Relative standard deviations of normalization factor k_{ij} between all datasets.

$\sigma(k_{ij})$ (%)	0	1	2	3	4	5	6	7
0	0.0	1.4	0.2	0.7	0.2	0.3	-	-
1	1.4	0.0	1.2	1.8	1.2	1.5	-	-
2	0.2	1.2	0.0	0.6	0.2	0.2	12.2	33.8
3	0.7	1.8	0.6	0.0	0.6	0.8	3.8	28.2
4	0.2	1.2	0.2	0.6	0.0	0.2	12.2	33.8
5	0.3	1.5	0.2	0.8	0.2	0.0	-	-
6	-	-	12.2	3.8	12.2	-	0.0	25.0
7	-	-	33.8	28.2	33.8	-	24.9	0.0

$P-values$ corresponding to the integral on $[\chi^2;\infty]$ range of the χ^2 distribution for $(n_{\hat{A}}-1)$ degrees of freedom. The usable data are only those which pass the χ^2 test and common mass set $\hat{A}$ are defined as following:

$$\{\hat{A}\} = \{\{A\}_i \cap \{A\}_j \setminus P-values\ (\chi^2_{ij}) > 1-CL\} \tag{2}$$

where $\{A\}_i$ (respectively $\{A\}_j$) are the measured masses of the i^{th} (respectively j^{th}) dataset and CL is the confidence level chosen at CL=0.995 for this work. The common mass number $n_{\hat{A}} = Card(\{\hat{A}\})$ is presented in table 1. Formally, the comparison of each dataset $\mathbf{N}_j(A)$ to the reference one $\mathbf{N}_i(A)$ generates $P-values$ lower than $1-CL = 0.005$ for all data set. So we introduce a cross-normalization factor $k_{i,j}$ to maximize the number of measurements in agreement considering all measurements as relative ones. This comparison is described by the $\mathbf{C_{ij}}$ vector following:

$$\mathbf{C_{ij}} = \mathbf{k_{ij}}.\mathbf{N_j} - \mathbf{N_i} \tag{3}$$

The normalization factors is obtained considering the minimum of the generalized χ^2_g over the $\{\hat{A}\}$ measurements in agreement:

$$\chi^2_g(k_{ij}) = \mathbf{C^T_{ij}}.\mathbf{Cov^{-1}}.\mathbf{C_{ij}} \tag{4}$$

where $\mathbf{Cov^{-1}}$ is the inverse covariance matrix associated to $\mathbf{C_{ij}}$. nevertheless, at this step without the covariance of the measurements, we consider that:

$$Cov(C_{ij}(A);C_{ij}(A')) = \left(Var(N_i(A) + k^2_{ij}.Var(N_j(A')\right).\delta_{AA'} \quad \forall A, A' \in \{\hat{A}\} \tag{5}$$

The standard deviations of the cross-normalization factors $k_{i,j}$ are presented in table 2. A discussion about usable data management is described in reference [15] [16].

2.2 Cross-correlations of usable data

The relative normalization of each j^{th} dataset, $N_j(A)$, to the (i^{th}) reference one, $N_i(A)$, is defined by as follows:

$$\mathbf{R^i_j} = \mathbf{k_{i,j}}.\mathbf{N_j} \tag{6}$$

According to the perturbation theory [17], the covariance of two normalized measurements $R^i_l(A)$ and $R^i_j(A')$ is developed in the appendix (see Sect. Apppendix). For this study, without explicit experimental covariance matrices, most of components of variance-covariance are considering null:

$$Cov(N_j(A); N_l(A')) = Var(N_j(A)).\delta_{AA'}.\delta_{jl} \quad \forall j, l$$

For n measurements of mass A [18], the mean normalized mass rate $\bar{R}(A)$ is equal to:

$$\bar{R}(A) = W_1.R^i_1(A) + ... + W_n.R^i_n(A) = \left(\sum_{j=1}^{n} W_j.R^i_j(A)\right) \tag{7}$$

For n measurements of mass A and m measurements of mass A' [18], the covariance of mean normalized mass rates is equal to (see Sect. Apppendix, eq. 23):

$$Cov(\bar{R}(A); \bar{R}(A')) = \sum_{l,j=1}^{n,m} W_l \, C_{lj} \, W_j \tag{8}$$

with here $C_{lj} = Cov(R^i_l(A); R^i_j(A'))$ for two different masses A and A', thus $(n \times m)$ terms. The weights are defined (see Sect. Apppendix, eq. 24):

$$W_{j=1,n} = \left(\sum_{l}^{n} C^{-1}_{lj}\right) / \left(\sum_{l,j}^{n,n} C^{-1}_{lj}\right) \quad and \quad W_{l=1,m} = \sum_{j}^{m} C^{-1}_{lj} / \sum_{l,j}^{m,m} C^{-1}_{lj} \tag{9}$$

with here $C_{lj} = Cov(R^i_l(A); R^i_j(A))$ for a same mass A with $(n \times n)$ covariance terms (or for a same mass A' with $(m \times m)$ covariance terms). Fig. 1 presents the cross-correlation of $R^i_l(A)$ data for two different reference sets. We note that the intensity of the correlation depends drastically of the choice of normalization and the uncertainty of the normalization factor $\sigma(k_{ij})$ (see table 2).

3 Ranking of solutions based on Shannon's entropy

According to the normalization of mass yields (see eq.1), the generalized perturbation theory [17] allows to describe the variance-covariance matrix associated to the evaluation of the mass yields:

$$\begin{aligned} \frac{Cov(Y(A); Y(A'))}{Y(A).Y(A')} &= \sum_{A''} S_{Y(A)\bar{R}(A'')} S_{Y(A')\bar{R}(A'')} \frac{Var(\bar{R}(A''))}{\bar{R}^2(A'')} \\ &+ \sum_{A'',A'''} S_{Y(A)\bar{R}(A'')}.S_{Y(A')\bar{R}(A''')}.\frac{Cov(\bar{R}(A''); \bar{R}(A'''))}{\bar{R}(A'').\bar{R}(A''')} \end{aligned} \tag{10}$$

with the sensitivity of mass yield Y(A) to mean mass rates $\bar{R}(A)$ and $\bar{R}(A')$:

$$S_{Y(A)\bar{R}(A)} = 1 - \frac{Y(A)}{\Omega} \quad and \quad S_{Y(A)\bar{R}(A')} = -\frac{Y(A')}{\Omega} \tag{11}$$

where Ω is the normalization factor, here $\Omega = 2$ (see eq. 1). Fig. 2 shows the results for two different reference sets and we remark that the structures of these correlations are strongly different. Thus, the result of the mass yields evaluation depends on the initial datasets but also the path of analysis. In order to discriminate all possible evaluations, the Shannon's entropy S_{Sh} is chosen as a useful criterion in order to assess the brewing of information [19]. It is given by the relation:

$$S_{Sh} = -\frac{1}{ln(2)} \sum_{k=1}^{n} P_i ln(P_i) \tag{12}$$

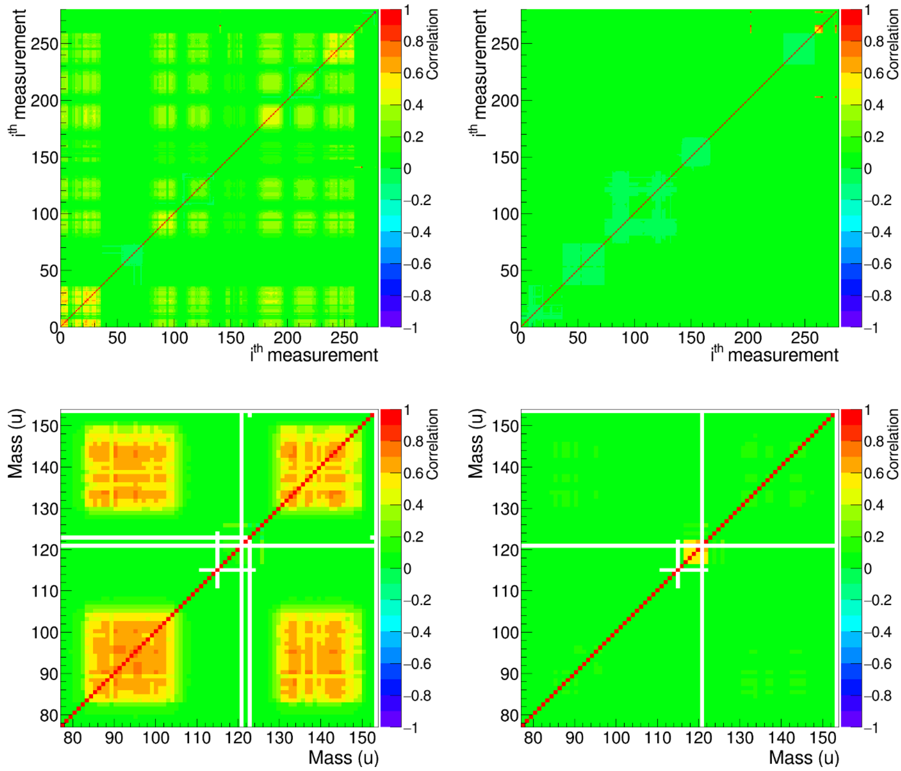

Figure 1. Preliminary results. (up, left) Cross-correlations for 8 datasets after normalization to the 3^{th} set; (up, right) the same data in reference to the 4^{th} set. These differences come from the uncertainties of the normalization factor $\sigma(k_{ij})$ which are not the same in both cases (see table 2). (down) Correlation of mean mass rate $Cov(\bar{R}(A); \bar{R}(A'))$ in reference to set 3 (left) and the set 4 (right).

Where n is the number of eigenvalues. We approximate the probability with the weight of each component of the eigenvalue decomposition to build a relative criterion. The weight of the information is provided according to the following equation:

$$P_i = \frac{EV_i}{tr(Corr)} \tag{13}$$

where $tr(Corr) = 77$ is the correlation matrix trace (in this study, 77 mass yields are evaluated).

4 Results and discussion

Results on pure experimental mass yields evaluation (modeless) are presented on Fig. 3. From all solutions, we note that the maximum of Shannon's entropy corresponds to the minimum of variances and correlations values. This results is consistent to the Cramer-Rao theorem which fixes the limits on minimal variances as the maximum of the Fischer's information. Shannon's entropy corresponds to another quantification of information of the analysis and we expect that the best searched solution corresponds to the minimum of variance-covariance and then

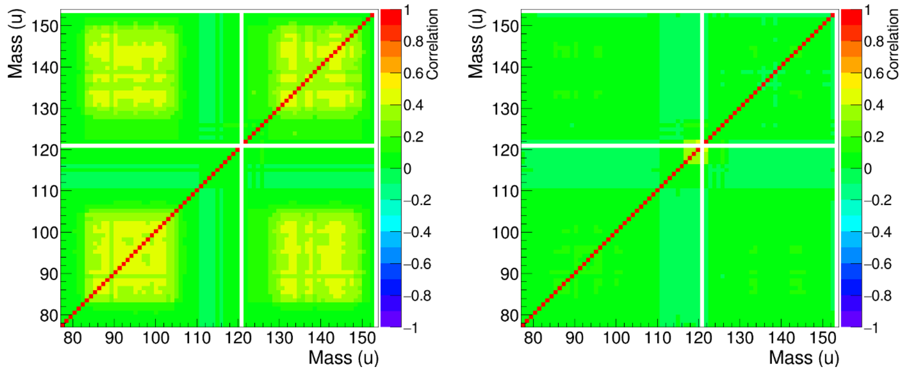

Figure 2. Preliminary results. (left) Mass yields correlations for 8 datasets after normalization to the 3^{th} set; (right) the same in reference to the 4^{th} set.

the maximum of information. In this work, experimental data consideration is crucial for the definition of the mass yields evaluation, the uncertainties and the correlations. The lack of experimental covariance could induce a lower estimation of evaluated mass yields uncertainties since the dealt information is overestimated. Correlations in the data limit the knowledge provided by a dataset. Then the perspective of this work is to build *a priori* experimental correlation matrix to fill the lacks in this analysis.

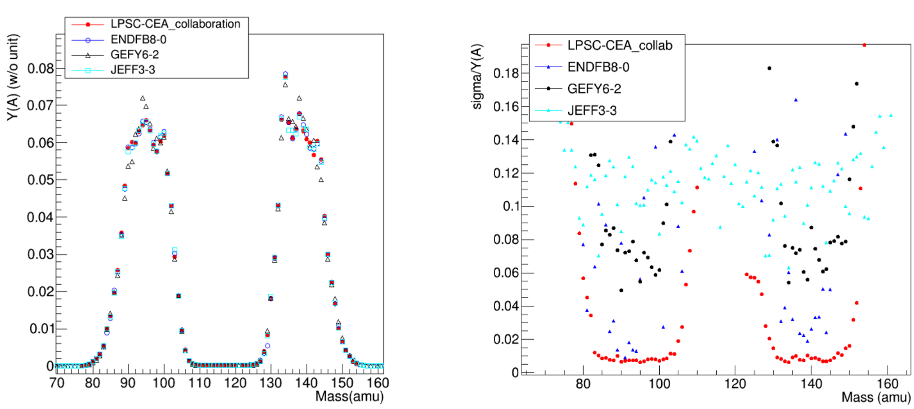

Figure 3. Preliminary results. (left) Mass yields evaluation proposed according to the maximum of Shannon's entropy (4^{th} reference set) in comparison to JEFF-3.3 [20] and ENDF/VIII.0 [21] libraries and the calculation from GEFY6-2 [3]. (right) Relative standard deviation of evaluated mass yields in comparison to those of evaluations or GEFY6-2.

Acknowledgment

This work was supported by IN2P3, by the University of Grenoble Alpes and by "le défi NEEDS" and the (WP4) SANDA european project.

Appendix

The relative normalization of each j^{th} dataset, $N_j(A)$, to the (i^{th}) reference one, $N_i(A)$, is defined as follows:

$$\mathbf{R_j^i} = \mathbf{k_{i,j}}.\mathbf{N_j} \tag{14}$$

Based on the minimum of χ^2, the variance of normalization factor k_{ij} is given by:

$$\begin{aligned} Var(k_{ij}) = & \frac{1}{K^2_{denom/j}} \sum_{\hat{A}} \left(\alpha^i_{ij}(\widehat{A})\right)^2 Var(N_i(\widehat{A})) + \frac{1}{K^4_{denom/j}} \sum_{\hat{A}} \left(\alpha^j_{ij}(\widehat{A})\right)^2 Var(N_j(\widehat{A})) \\ & + \frac{2}{K^2_{denom/j}} \sum_{\hat{A}} \sum_{\widehat{A'}>\widehat{A}} \left(\alpha^i_{ij}(\widehat{A})\right)\left(\alpha^i_{ij}(\widehat{A'})\right) Cov(N_i(\widehat{A}), N_i(\widehat{A})) \\ & + \frac{2}{K^4_{denom/j}} \sum_{\widehat{A}} \sum_{\widehat{A'}>\widehat{A}} \left(\alpha^j_{ij}(\widehat{A})\right)\left(\alpha^j_{ij}(\widehat{A'})\right) Cov(N_j(\widehat{A}), N_j(\widehat{A'}) \end{aligned} \tag{15}$$

According to the perturbation theory, the covariance of two normalized measurements $R^i_l(A)$ and $R^i_j(A')$ is described by the following equation [17]:

$$\begin{aligned} Cov(R^i_l(A), R^i_j(A')) = & \frac{N_l(A)N_j(A')}{K_{dénom/j}K_{dénom/l}} \sum_{\widehat{A}} \sum_{\widehat{A'}} \alpha^i_{ij}(\widehat{A})\alpha^i_{il}(\widehat{A'}) Cov(N_i(\widehat{A}), N_i(\widehat{A'})) \\ & + \frac{N_l(A)N_j(A')}{K_{dénom/j}K^2_{dénom/l}} \sum_{\widehat{A}} \sum_{\widehat{A'}} \alpha^i_{ij}(\widehat{A})\alpha^l_{il}(\widehat{A'}) Cov(N_i(\widehat{A}), N_l(\widehat{A'})) \\ & + \frac{N_l(A)N_j(A')}{K^2_{dénom/j}K_{dénom/l}} \sum_{\widehat{A}} \sum_{\widehat{A'}} \alpha^i_{il}(\widehat{A'}\alpha^j_{ij}(\widehat{A}) Cov(N_j(\widehat{A}), N_i(\widehat{A'})) \\ & + \frac{N_l(A)N_j(A')}{K^2_{denom/l}K^2_{denom/j}} \sum_{\widehat{A}} \sum_{\widehat{A'}} \alpha^j_{ij}(\widehat{A})\alpha^l_{il}(\widehat{A'}) Cov(N_j(\widehat{A}), N_l(\widehat{A'})) \\ & + \frac{k_{il}N_j(A)}{K_{dénom/j}} \left(\sum_{\widehat{A'}} \alpha^i_{ij}(\widehat{A'}) Cov(N_l(A), N_i(\widehat{A'})) + \frac{1}{K_{dénom/j}} \sum_{\widehat{A'}} \alpha^j_{ij}(\widehat{A'}) Cov(N_l(A), N_j(\widehat{A'})) \right) \\ & + \frac{k_{ij}N_l(A)}{K_{dénom/l}} \left(\sum_{\widehat{A}} \alpha^i_{il}(\widehat{A}) Cov(N_j(A'), N_i(\widehat{A})) + \frac{1}{K_{dénom/l}} \sum_{\hat{A}} \alpha^j_{il}(\widehat{A}) Cov(N_j(A'), N_l(\widehat{A})) \right) \\ & + k_{il}k_{ij} Cov(N_l(A), N_j(A')) \end{aligned} \tag{16}$$

with:

$$\alpha^i_{ij} = \frac{N_j(\hat{A})}{\sigma^2_{ij}} \qquad \alpha^j_{ij} = \frac{(K_{denom/j}N_i(\hat{A}) - 2K_{num/j}N_j(\hat{A}))}{\sigma^2_{ij}} \tag{17}$$

$$K_{num/j} = \sum_{\hat{A}} \frac{N_i(\widehat{A})N_j(\widehat{A})}{\sigma^2_{ij}} \qquad K_{denom/j} = \sum_{\hat{A}} \frac{N^2_j(\hat{A})}{\sigma^2_{ij}} \tag{18}$$

For this study, without explicit experimental covariance matrices, most of components of variance-covariance are considering null:

$$Cov(N_j(A); N_l(A')) = Var(N_j(A)).\delta_{AA'}.\delta_{jl} \quad \forall j, l \tag{19}$$

For n measurements of mass A, the mean normalized mass rate $\bar{R}(A)$ is equal to [18]:

$$\bar{R}(A) = W_1.R_1^i(A) + ... + W_n.R_n^i(A) = \left(\sum_{j=1}^{n} W_j.R_j^i(A)\right) \tag{20}$$

with:

$$\begin{aligned} W_{j=1,n} &= \frac{C_{1j}^{-1} + ... + C_{nj}^{-1}}{C_{11}^{-1} + C_{12}^{-1} + ... + C_{1n}^{-1} + C_{21}^{-1} + ... + C_{2n}^{-1} + ... + C_{n1}^{-1} + ... + C_{nn}^{-1}} \\ &= \left(\sum_{l}^{n} C_{lj}^{-1}\right) / \left(\sum_{l,j}^{n,n} C_{lj}^{-1}\right) \end{aligned} \tag{21}$$

with $C_{lj} = Cov(R_l^i(A); R_j^i(A))$ for a same mass A and its variance is given by the following equation:

$$Var(\bar{R}(A)) = 1/\left(\sum_{lj}^{n,n} C_{lj}^{-1}\right) \tag{22}$$

For n measurements of mass A and m measurements of mass A', the covariance of mean normalized mass rates is equal to:

$$Cov(\bar{R}(A); \bar{R}(A')) = \sum_{l,j=1}^{n,m} W_l\, C_{lj}\, W_j \tag{23}$$

with here $C_{lj} = Cov(R_l^i(A); R_j^i(A'))$ for two different masses A and A', thus $(n \times m)$ terms; and the weights:

$$W_{j=1,n} = \left(\sum_{l}^{n} C_{lj}^{-1}\right) / \left(\sum_{l,j}^{n,n} C_{lj}^{-1}\right) \quad and \quad W_{l=1,m} = \sum_{j}^{m} C_{lj}^{-1} / \sum_{l,j}^{m,m} C_{lj}^{-1} \tag{24}$$

with here $C_{lj} = Cov(R_l^i(A); R_j^i(A))$ for a same mass A with $(n \times n)$ covariance terms (or for same mass A' with $(m \times m)$ covariance terms)

References

[1] L. Fiorito et al. Annals of Nuclear Energy, **69**, (2014) 331-343.
[2] A. Chebboubi et al., EPJ Web of Conferences, **146**, (2017) 04021.
[3] K. H. Schmidt et al., Nucl. Data Sheets, **131**, (2016) 107-221.
[4] N. Terranova et al., Nucl. Data Sheets **95**, (2015) 225–230.
[5] D. Rochman et al., Annals of Nuclear Energy, **131**, (2016) 125-134.
[6] O. Leray et al.,EPJ Web of Conferences, **146**, (2017) 09023.
[7] EXFOR database, https://www-nds.iaea.org/exfor/exfor.htm
[8] W. J. Maeck et al., Allied Chem. Corp., Idaho Chem. Programs, **1142**, (1978) 09
[9] G. Diiorio and B. W. Wehring, Nuclear Instruments and Methods, **147**, (1977) 487 - 499.
[10] H. Thierens et al., Nuclear Instruments and Methods, **134**, (1976) 299 - 308.
[11] A. Bail, PhD Thesis, Université Sciences et Technologies - Bordeaux I, (2009).
[12] M.P. Tsoukatos et al., Dissertation Abstracts B (Sciences), volume 28, page 2759, 1968

[13] K.J. Rosman, Canadian Journal of Physics, volume 61, page 1490, 1983

[14] C.K. Mathiews, Physical Review, Part C, Nuclear Physics, volume 15, page 344, 1977/01

[15] B. Voirin, et al., EPJ N - Nuclear Sciences & Technologies, EDP Sciences, 2018, 4, pp.26

[16] S. Julien-Laferrière, et al., EPJ N - Nuclear Sciences & Technologies, EDP Sciences, 2018, 4, pp.25

[17] D.L. Smith, *Probability, statistics, and data uncertainties in nuclear science and technology, vol.4*, (OECD,Paris, 1991) 269p

[18] M. Schmelling, Phys. Scr. 51 676, 1995

[19] B. Diu, C. Guthmann, B. Roulet, D. Lederer, *Physique statistique* (Hermann, 1996) ISBN: 2-7056-6065-8

[20] The JEFF-3.1 nuclear data library, OECD, NEA, JEFF Report 21, ISBN 92-64-02314-3, (2006).

[21] M. B. Chadwick, et al., Nucl. Data Sheets **112** (2011) 2887.