What is the proper evaluation method: Some basic considerations

Helmut Leeb, Georg Schnabel, Thomas Srdinko
Atominstitut, Technische Universität Wien
Vienna, Austria

Abstract

Recent developments and applications demand for an extension of the energy range and the inclusion of reliable uncertainty information in nuclear data libraries. Due to the scarcity of neutron-induced reaction data beyond 20 MeV the extension of the energy range up to at least 150 MeV is not trivial because the corresponding nuclear data evaluations depend heavily on nuclear models and proper evaluation methods are still under discussion. Restricting to evaluation techniques based on Bayesian statistics the influence of the a priori knowledge on the final result of the evaluation is considered. The study clearly indicates the need to account properly for the deficiencies of the nuclear model. Concerning the covariance matrices it is argued that they depend not only on the model, but also on the method of generation and an additional consent is required for the comparison of different evaluations of the same data sets.

Introduction

The availability of evaluated nuclear data of high quality is an important prerequisite for the development of novel nuclear technologies, construction of nuclear installations, determination of safety measures as well as medical and industrial applications. Hence there is a world wide effort to obtain consolidated data sets (i.e. cross-sections, spectra, etc.) which represent our best knowledge of observables relevant for the user community. For many key observables accurate and complete measurements have been performed or are planned. Including the new measurements an update of the consolidated data sets must be performed taking consistently into account all available previous data sets as well as the *a priori* knowledge usually expressed in terms of nuclear models and theory constraints. The consolidation process is usually denoted as nuclear data evaluation and provides a consistent set of values of observables as well as uncertainty estimates. Well known examples of such evaluations are the set of consistent fundamental constants (Mohr, Taylor and Newell, 2012) and the consolidated particle data (Beringer, *et al.*, 2012). In principle such an evaluation step is required for any scientific conclusion from experiment, but frequently one considers dedicated experiments with an almost unique interpretation where the evaluation step is trivial.

In this contribution we focus on nuclear data evaluation, which provides an extended set of nuclear cross-section data and spectra consistent with available experimental data and satisfying all basic principles (e.g. sum rules, unitarity, …). As mentioned above there is a wide community of users in industry, medicine and science. Fuelled by recent developments in nuclear technology and accelerator applications there is a demand for an increased energy range (at least up to 150 MeV) and the inclusion of uncertainty information in nuclear data libraries. The latter is also driven by questions of optimisation and consequently by economics. The extension of the energy range is not trivial because of the scarcity of experimental data sets at incident neutron energies beyond 20 MeV. Hence a corresponding evaluation relies heavily on the prior and thus on the nuclear model used. Despite the world-wide effort to develop proper methods for evaluations which rely heavily on nuclear models the problem is still not solved. A recent status review of the developed evaluation procedures is given in Bauge, *et al.* (2010). Among these the Bayesian techniques are most promising because they properly account for *a priori* knowledge in a natural way. Though the procedures are well defined, the meaning of the resulting uncertainty information in terms of probability is not fully clarified. Thus no established criteria are available for comparison and quality assessment at present.

In the following we study the properties of Bayesian evaluation techniques considering a specific example in order to obtain a deeper insight in the properties of the evaluation with regard to mean values and uncertainties. Therefore we briefly revisit the basics of Bayesian evaluation techniques in the next section. Specifically, for the example of a simple evaluation of some neutron-induced reaction channels of ^{181}Ta a careful analysis of mean values and covariance matrices is performed. This study clearly indicates the need to account for the deficiencies of models in nuclear data evaluations. In the subsequent section, *Open problems of nuclear data evaluation*, the impact of model defects as defined in Koning, Hilaire and Duijvestijn (2008) is discussed and open problems in covariance matrices are addressed. In the final section concluding remarks are given.

Bayesian evaluation techniques

Concept

It is the primary goal of nuclear data evaluation to provide the best knowledge of a given set of observables $\boldsymbol{\sigma}$ based on available measurements and *a priori* knowledge. The latter is of mathematical and/or physics nature and is an inherent feature of the involved nuclear models depending on a set of parameters $\mathbf{x}$.

Bayesian statistics is generally accepted as the proper means for a consistent combination of measured data and *a priori* knowledge. It is based on two simple principles, i.e.:

The sum rule: $$p(\mathbf{x}|M)+p(\bar{\mathbf{x}}|M)=0 \quad (1)$$

The product rule: $$p(\mathbf{x}|\sigma M)\,p(\sigma|M)=p(\sigma|\mathbf{x}M)\,p(\mathbf{x}|M) \quad (2)$$

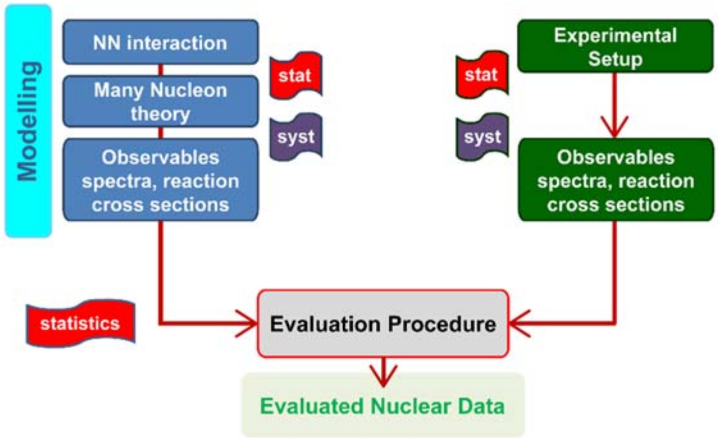
where $p(a|B)$ denotes a conditional probability density function (pdf) that an event a takes place if the condition B is true. Here, M refers to the specific nuclear model chosen. The basis of nuclear data evaluation is Bayes theorem:

$$p(\mathbf{x}|\sigma M)=C\cdot p(\sigma|\mathbf{x}M)\,p(\mathbf{x}|M) \quad (3)$$

which is a consequence of Eq. (2). It combines *a priori* knowledge given in terms of the prior pdf $p(\mathbf{x}|M)$ with experimental knowledge via the likelihood function $p(\sigma|\mathbf{x}M)$. The quantity C is a normalisation constant. Eq. (3) is the basic relation of Bayesian evaluation methods [see e.g. Bauge, *et al.* (2010)].

In the following we will use the so-called "Full Bayesian Evaluation Technique" (Leeb, Neudecker and Srdinko, 2008) for our considerations. As shown in Figure 1, a Bayesian evaluation procedure has two types of input, i.e. *i)* the experimental data and their uncertainties; *ii)* the *a priori* knowledge which is given in terms of nuclear models. Both informations are associated with probability density functions and determine via Eq. (3) the *a posteriori* pdf which in turn gives the mean values and the covariance matrices in the evaluation step.

Figure 1: General scheme of an evaluation procedure based on Bayesian statistics



Both the experimental data as well as the prior are subject to statistical uncertainties and therefore the evaluation in terms of Bayesian statistics is best suited. In addition systematic errors occur which are of non-statistical nature and require a special treatment (Leeb, Neudecker and Srdinko, 2008).

Properties of the evaluation

In order to reveal the generic properties of a Bayesian evaluation technique we study a simplified evaluation restricted to the total and differential elastic scattering cross-sections of neutrons incident on ^{181}Ta. The prior is generated by default calculations with TALYS-1.4 (Koning, Hilaire and Duijvestijn, 2008) and depends, apart from the compound-elastic contribution, only on the optical potential. The statistical uncertainties are generated by random variation of the parameters of the optical model. Actually a Monte Carlo technique is applied assuming a uniform distribution between physically reasonable defined boundaries. At first glance this seems to be an *ad hoc* procedure and a maximum entropy based method (Leeb and Pigni, 2006) should be more appropriate. However, it turned out that the prior is mainly affected by the selected boundaries of the parameters, while the probability density functions of the parameters are of minor importance. Experimental total cross-sections (Finlay, *et al.*, 1993) and differential elastic cross-sections (Smith, 2005) with properly estimated covariance matrices have been used for the evaluation. For the evaluation procedure we choose a reasonable mesh in energy $\{E_i\}$ and represent the angle integrated cross-sections $\sigma_c(E)$ with the help of spline functions. For the differential cross-section we choose the representation:

$$\frac{d\sigma}{d\Omega} = \sum_{\lambda} \alpha_{\lambda}(E) P_{\lambda}(\cos\vartheta) \tag{4}$$

where $P_{\lambda}(\cos\vartheta)$ are Legendre polynomials and $\alpha_{\gamma}(E)$ are coefficient functions which again are described by spline functions using the same energy mesh. It is important to remark that in this representation the values of $\sigma_c(E)$ and $\alpha_{\gamma}(E)$ at the mesh points $\{E_i\}$ act as the parameters of our "model" in the evaluation. Since a multivariate normal distribution is imposed for the cross-sections at the mesh points, sum rules are exactly maintained there. For a sufficiently dense grid this will also be true in good approximation in between the mesh points. The Bayesian evaluation procedure has been performed in its linearised version as outlined in Leeb, Neudecker and Srdinko (2008).

The results of this simple evaluation are displayed in Figures 2 and 3. In Figure 2 the prior of the differential elastic cross-section of n-^{181}Ta and the corresponding evaluated data at $E = 4.51$ MeV are shown. It is evident from Figure 2 that the evaluation

Figure 2: Comparison of experimental differential elastic n-^{181}Ta cross-section of Smith (2005) at E = 4.51 MeV with (a) the prior cross-section and its uncertainty band and (b) the evaluated differential cross-section

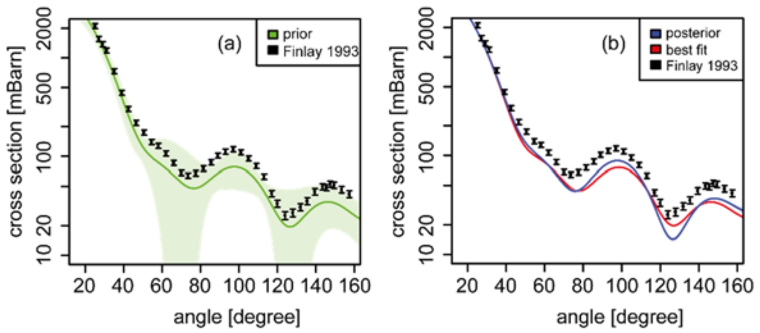
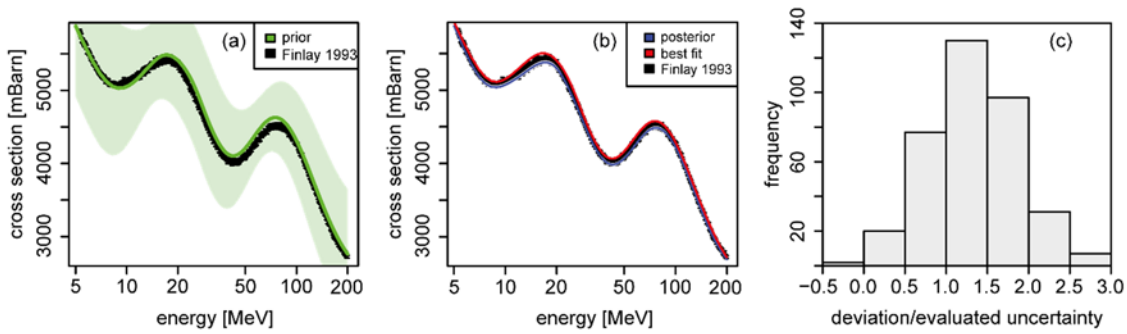


Figure 3: Experimental total n-^{181}Ta cross-section (Finlay, *et al.*, 1993) compared with (a) the prior and (b) the evaluated cross-section; (c) deviation: evaluated from experimental one



underestimates the cross-section. The same is also found at other energies. There arises the question whether this behaviour is a specific feature of angle-differential data. Therefore we also took a closer look at the total cross-section shown in Figure 3. At first glance the evaluated total cross-section agrees extremely well with the experimental values. A more quantitative analysis given in the histogram, part (c) of Figure 3, reveals again a small but systematic underestimation of the total cross-section by the evaluation.

Hence the problem of systematic deviation of the evaluated cross-section from the mean of experimental data appears to be a general one and is not a specific feature of certain observables. In addition for both observables the evaluation leads to unrealistically small error bands. These observations jeopardise the usefulness of any evaluation and require a more careful analysis. With regard to a compact presentation we restrict the analysis here to the total cross-section.

In a first step we performed a principal component analysis of the prior covariance matrix of total cross-section uncertainties. The ordered eigenvalues, i.e. the variances in direction of the principal axes, are shown in Figure 4. In Figure 5 the 12 most important normalised eigenvectors $\phi_i(E)$ are displayed which reproduce the covariance matrix of total cross-section uncertainties better than 0.3%. From this consideration it is obvious that the evaluated total cross-section must be a superposition of the eigenvectors. In particular the mean value of the total cross-section of the prior is given by:

$$\sigma_{tot}(E) = \sum_{i=1}^{K} a_i \phi_i(E) \qquad (5)$$

where K is the number of mesh points in energy. In Table 1 the coefficients a_i and their standard deviation are listed. For a model prior which gives a fair description of the experimental data the first 16 eigenstates will constitute the dominant contribution and uncertainty bands comparable to experimental capabilities will emerge. If the model prior significantly deviates from the experimental data the evaluated cross-section will be mainly composed of eigenstates with small variances. Therefore the covariance matrix of the evaluated total cross-section uncertainties will have unrealistically small variances. Consequently, these unrealistically small error bands reflect the deficiencies of the model and are not an indication of excellent knowledge of the observables.

The problem can be easily illustrated by a schematic example. Let us assume a model which allows constant energy-independent cross-sections denoted with $\mathbf{x}_{prior}$. The corresponding prior covariance matrix of cross-section uncertainties A_{prior} is of rank 1 with one eigenvector associated with a non-zero eigenvalue $n\delta^2_{mod}$, where n is the number of mesh points and δ^2_{mod} is the variance at every mesh point. Furthermore we assume a set of experimental data $\mathbf{y}_{exp}$ which exhibit an energy dependence and are

Figure 4: Eigenvalues of the prior covariance matrix (see text)

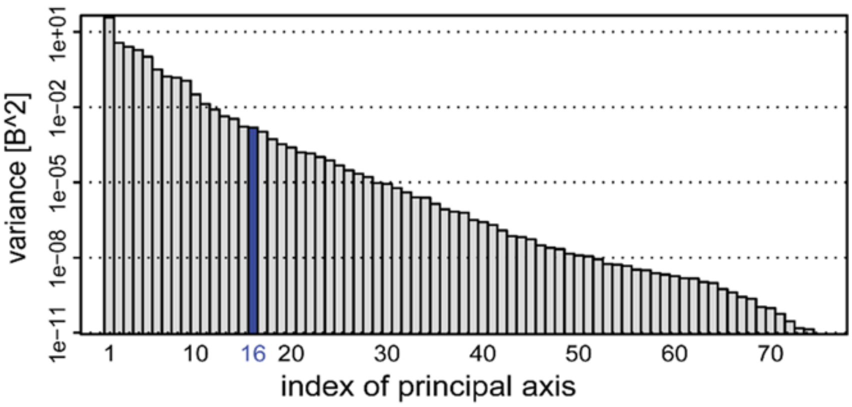


Figure 5: The first 12 normalised eigenvectors of the prior covariance matrix of total cross-section uncertainties

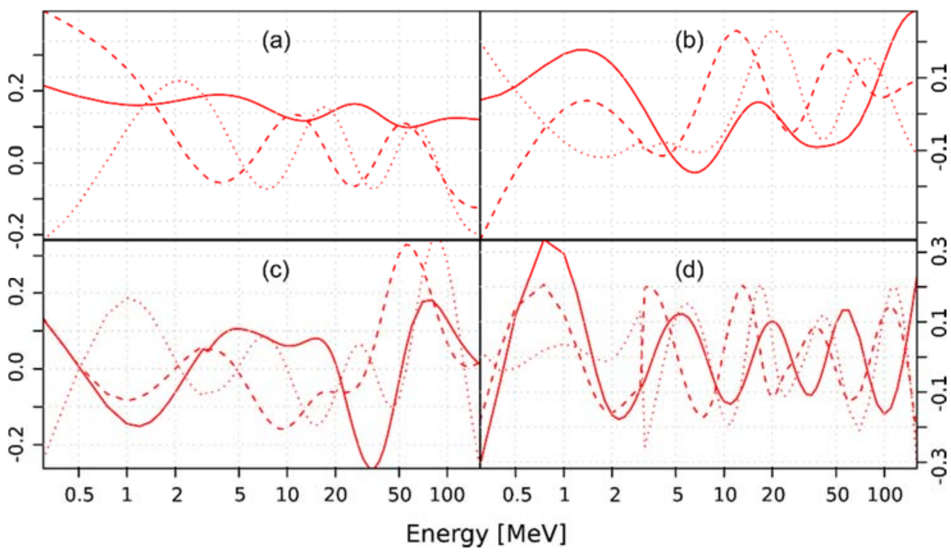


Table 1: Coefficients a_i and standard deviation Δa_i for the first 12 eigenvectors

"EV No." (eigenvector number), "Line type" and "Part" (a), (b), (c) and (d) refer to Figure 5

Line type	Part (a)			Part (b)		
	EV No.	a_i (barn)	Δa_i (barn)	EV No.	a_i (barn)	Δa_i (barn)
Solid	1	48.669	6.122	4	1.415	1.377
Dashed	2	6.147	1.928	5	5.077	1.012
Dotted	3	1.378	1.599	6	0.752	0.564
Line type	Part (c)			Part (d)		
	EV No.	a_i (barn)	Δa_i (barn)	EV No.	a_i (barn)	Δa_i (barn)
Solid	7	1.112	0.411	10	0.165	0.180
Dashed	8	0.773	0.389	11	0.107	0.116
Dotted	9	0.665	0.336	12	0.041	0.090

given at the same mesh points as the model. The experimental covariance matrix B_{ij} may include a statistical error Δ^2_{stat} and an overall normalisation error Δ^2_{syst} affecting every experimental value:

$$B_{ij} = \Delta^2_{stat}\delta_{ij} + \Delta^2_{stat} \tag{6}$$

The corresponding sensitivity matrix is the identity. Thus the linearised update can be performed in closed form and yields the *a posteriori* covariance matrix:

$$A_{post} = \frac{1}{1+\dfrac{n\delta^2_{\mathrm{mod}}}{\Delta^2_{stat}+n\Delta^2_{syst}}} A_{prior} = \frac{\delta^2_{\mathrm{mod}}}{1+\dfrac{n\delta^2_{\mathrm{mod}}}{\Delta^2_{stat}+n\Delta^2_{syst}}} \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & & 1 & 1 \\ \vdots & & \ddots & & \vdots \\ 1 & 1 & & 1 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \tag{7}$$

This toy model shows several characteristic features. The posterior covariance matrix is completely determined by the uncertainty estimates of the model and the experiment. The mean value of the prior and the experimental value do not enter into the posterior covariance matrix. For vanishing systematic errors and $\delta_{\mathrm{mod}} >> \Delta_{stat}$ all elements of the posterior covariance matrix are Δ^2_{stat}/n and thus much smaller than the experimental uncertainties. For vanishing statistical error and $\delta_{\mathrm{mod}} >> \Delta_{syst}$ the evaluated uncertainty is approximately given by Δ_{syst} of the experiment. For an increasing number of experimental data the influence of the statistical error on the evaluated uncertainties vanishes.

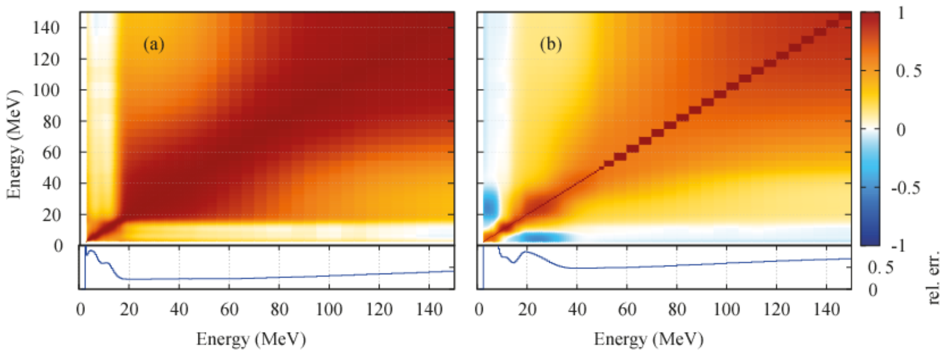
Open problems of nuclear data evaluation

Accounting for model deficiencies

The considerations in the previous section clearly demonstrate the importance to account for deficiencies of the nuclear model in the evaluation procedure. In nuclear physics the occurrence of deficient descriptions is obvious because we deal with a many-body problem which cannot be solved in its full complexity. Hence nuclear models have been set up which describe part of the effects sufficiently well, while other aspects are not reproduced at all. A first systematic formulation of a covariance matrix associated with model defects was given in Leeb, Neudecker and Srdinko (2008). An alternative procedure for model defects is used by Trkov, *et al.* (2011). The method of Leeb, Neudecker and Srdinko (2008) is based on a systematic comparison of experimental reaction data of neighbouring nuclei with model calculations. In this procedure it is inherently assumed that the model is equally well suited for these nuclei and reliable estimates of the deviations can be obtained. In Figure 6 a typical prior covariance matrix without and with model defects is shown for the ^{55}Mn(n,p) reaction cross-section uncertainties. The most important effect of model deficiencies on covariance matrices is obviously the reduction of correlations. Having this in mind we have repeated the simplified evaluation of n-^{181}Ta with a non-correlated prior keeping only the diagonal elements. As expected we found in this case agreement of the evaluated cross-sections with the experimental data and the uncertainty bands remained in the order of the systematic errors of the experiments. Recently a more sophisticated study (Neudecker, Capote and Leeb, 2013) on ^{55}Mn(n,2n) reactions was presented which clearly demonstrates that the inclusion of model defects is essential to obtain reasonable error bars of the evaluated cross-sections. Apart from the success of available formulations to account for model deficiencies their use in complete evaluations is still questionable because violations of sum rules and unitarity may occur.

Figure 6: Correlation matrix and relative errors of the ^{55}Mn(n,p) cross-section

(a) Parameter uncertainties only, (b) model deficiencies included



Covariance matrices associated with modelling

The covariance matrices for experimental cross-section uncertainties are well defined by means of statistics. However the situation is more intriguing if we determine the covariance matrix starting from a nuclear model. This is easily seen, if we assume linear error propagation in the vicinity of the best set of parameters **p**:

$$A_{x,x'} = \left\langle \Delta\sigma(x)\Delta\sigma(x')\right\rangle = \sum_{i,j=1}^{N} \frac{\partial\sigma(x,\mathbf{p})}{\partial p_i}\left\langle \Delta p_i \Delta p_j\right\rangle \frac{\partial\sigma(x,\mathbf{p})}{\partial p_j} \tag{8}$$

Here, it is assumed that the nuclear model depends on N parameters $p_1,p_2,\ldots,p_N$ and x and x' are values of the independent variable, e.g. the energy. At present the priors are usually generated by uncorrelated variation of the parameters p_i. However, one may use any other correlation of the parameters and obtain a different covariance matrix. This situation is not satisfactory with regard to comparison of evaluations and in particular for applications. Therefore additional criteria would be useful to classify evaluations.

The previous question is not trivial and depends on the quality of the model with regard to physics. Let us assume that we know for a given observable $y(x)$ the true physics description $y(x) = f(x,\mathbf{p})$, where **p** are N parameters and x is the independent variable, e.g. the energy. In general the physics description is very involved and therefore we prefer to use a model description $y(x) = g(x,\mathbf{q})$ which depends on M parameters $\mathbf{q}$ and is easier to evaluate. On a given mesh $\{x_i\}$ the model yields the same values of the observable as the complete physics description. Introducing parameter covariance matrices and associated sensitivity matrices:

$$F_{i,j} = \langle \Delta p_i \Delta p_j \rangle,\ S_{i,j} = \frac{\partial f(x_i,\mathbf{p})}{\partial p_j} \text{ and } G_{i,j} = \langle \Delta q_i \Delta q_j \rangle,\ R_{i,j} = \frac{\partial g(x_i,\mathbf{q})}{\partial q_j} \qquad (9)$$

one obtains the covariance $A^{\mathrm{phy}} = S\,F\,S^{\mathrm{T}}$ and $A^{\mathrm{mod}} = R\,G\,R^{\mathrm{T}}$ for the physics and the model description, respectively. Requiring that both covariance matrices are equal, $A^{\mathrm{phy}} = A^{\mathrm{mod}}$, and assuming a diagonal matrix F implies different correlations of the model parameters. In particular if the model space is larger than the physics dependences the matrix G must contain dependences. In the case that the model space is too small, the model parameter covariance matrix G is a projection of F. Hence the quality of the model has also an impact on the proper covariance matrix.

Conclusions

It has been the primary goal of this contribution to reveal basic features of a Bayesian evaluation process. In particular we have shown in an example that the evaluation based on the generation of prior covariance matrices via parameter variation may yield questionable results with respect to mean values and uncertainties. This is also true if within the model a perfect description of the experimental data is in principle possible but unlikely. Especially the variances of the evaluation may result unrealistically small. This is not a failure of the Bayesian approach, but reflects the fact that the evaluated covariance matrices are always obtained for a specific model as indicated by the symbol M in Eq. (3). Our considerations clearly indicate the necessity to account for model deficiencies in the evaluation process. These model deficiencies lead to reduced correlations of the covariance matrices. The use of available formulations of model deficiencies is very promising, but still has the drawback that basic features, e.g. sum rules, are violated in complete evaluations. Therefore work is in progress to find better formulations which solve this problem.

In addition we pointed out that the generation of covariance matrices from nuclear models are not uniquely defined. Additional criteria are needed to classify evaluations with respect to their uncertainties. It is also argued that the quality of the model has to be taken into account in these considerations.

In summary for the quality of an evaluation strongly based on nuclear models it is important to use the best available physics description. Thus model deficiencies will become small, but are still required to accomplish reliable evaluations. In fact nuclear data evaluation is a challenge for both experimental and theory research. In this context the CIELO project is an important initiative which will provide not only better evaluations for selected nuclei, but also deeper insight into their physics.

Acknowledgements

Work has been supported by EURATOM via the specific grant of the Fusion for Energy Partnership Agreement F4E-FPA-168. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

Bauge, E., *et al.* (2011), *Covariance Data in the Fast Neutron Region*, International Evaluation Co-operation, Vol. 24, NEA/NSC/WPEC/DOC/(2010)427, OECD/NEA, Paris.

Beringer, J. *et al.* (2012), "Review of Particle Physics", *Phys. Rev. D*, 86, 010001.

Finlay, R.W., *et al.* (1993), "Neutron Total Cross Sections at Intermediate Energies", *Phys. Rev. C*, 47, pp. 237-247.

Koning, A.J., S. Hilaire and M.C. Duijvestijn (2008), "TALYS-1.0", *Proceedings of the International Conference on Nuclear Data for Science and Technology*, Nice, France, 22-27 April 2007, O. Bersillon, *et al.* (eds.), EDP Sciences, pp. 211-214.

Leeb, H., D. Neudecker and T. Srdinko (2008), "Consistent Procedure for Nuclear Data Evaluation Based on Modelling", *Nucl. Data Sheets*, 109, pp. 2762-2767.

Leeb, H. and M.T. Pigni (2006), "Basic Statistics and Consistent Covariances for Nuclear Data Files", *Proceedings of the Workshop on Perspectives of Nuclear Data in the Next Decade*, Bruyères-le-Châtel, France, E. Bauge (ed.), NEA No. 6121, OECD/NEA, Paris, pp. 235-242.

Mohr, P.J., B.N. Taylor and D.B. Newell (2012), "CODATA Recommended Values of the Fundamental Physical Constants: 2010", *Rev. Mod. Phys.*, 84, pp. 1527-1605.

Neudecker, D., R. Capote and H. Leeb (2013), "Impact of Model Defect and Experimental Uncertainties on Evaluated Output", *Nucl. Instr. Meth. A*, 723, pp. 163-172.

Smith, A.B. (2005), "Fast Neutrons Incident on Rotors: Tantalum", *Annals of Nuclear Energy*, 32, pp. 1926-1952.

Trkov, A., *et al.* (2011), "Covariances of Evaluated Nuclear Cross Section Data for ^{232}Th, ^{180,182,183,184,186}W and ^{55}Mn", *Nuclear Data Sheets*, 112, pp. 3098-3119.