

Nuclear data assimilation, scientific basis and current status

Evgeny Ivanov^{1,*}, Cyrille De Saint-Jean², and Vladimir Sobes³

¹ Institute for Radioprotection and Nuclear Safety, Fontenay-aux-Roses, France

² Alternative Energies and Atomic Energy Commission, CEA DAM DIF, F-91297 Arpajon, France

³ The University of Tennessee, Knoxville, USA

Received: 9 November 2020 / Received in final form: 24 March 2021 / Accepted: 9 April 2021

Abstract. The use of Data Assimilation methodologies, known also as a data adjustment, liaises the results of theoretical and experimental studies improving an accuracy of simulation models and giving a confidence to designers and regulation bodies. From the mathematical point of view, it approaches an optimized fit to experimental data revealing unknown causes by known consequences that would be crucial for data calibration and validation. Data assimilation adds value in a ND evaluation process, adjusting nuclear data to particular application providing so-called optimized design-oriented library, calibrating nuclear data involving IEs since all theories and differential experiments provide the only relative values, and providing an evidence-based background for validation of Nuclear data libraries substantiating the UQ process. Similarly, it valorizes experimental data and the experiments, as such involving them in a scientific turnover extracting essential information inherently contained in legacy and newly set up experiments, and prioritizing dedicated basic experimental programs. Given that a number of popular algorithms, including deterministic like Generalized Linear Least Square methodology and stochastic ones like Backward and Hierarchic or Total Monte-Carlo, Hierarchic Monte-Carlo, etc., being different in terms of particular numerical formalism are, though, commonly grounded on the Bayesian theoretical basis. They demonstrated sufficient maturity, providing optimized design-oriented data libraries or evidence-based backgrounds for a science-driven validation of general-purpose libraries in a wide range of practical applications.

1 Introduction

The first practical use of Data Assimilation (DA) in the nuclear engineering started in the sixties to take a maximum of benefits from rare – that time – experimental data, developing nuclear reactor design concepts and improving problem-oriented nuclear data libraries [1–5].

From the very beginning it involved the deterministic algorithms, such as a Generalized Linear Least Square methodology (GLLSM) associated that time with Ordinary and Generalized Perturbation Theory (OPT and GPT) [1,2,6,7]. Further trend was traced in a progressive growth of DA fidelity supported by an increasing of computational capacities and advances in mathematical and computational physics [8–17].

Nowadays, available and affordable highly performant computational facilities and high-fidelity or, even, precise

code systems allow stepping toward fully stochastic non-intrusive¹ adjustment [12–16].

Despite on the continuous evolution of the DA algorithms the methodology always remains to be a kind of Bayesian-based technique [1,5,7,8].

It should be noted that DA is always present in the Nuclear Data (ND) evaluation process because all, without any exemption, data libraries have to be somehow calibration² on the objective experimentally measured invariants. DA helps, in this case, if such invariants were not measured directly but inferred from the measurements [2,13].

In any applications the mathematical models and data libraries to become suitable for the adjustment should be somehow parametrized using either Reduced Order Models (ROM) or variables inherent to nuclear reactions simulations [1,13,15,16].

¹ The term “non-intrusive” means that one could apply high-fidelity analytical tools without simplifications like sensitivity calculations or so on.

² As any physical theory.

*e-mail: evgeny.ivanov@irsrn.fr

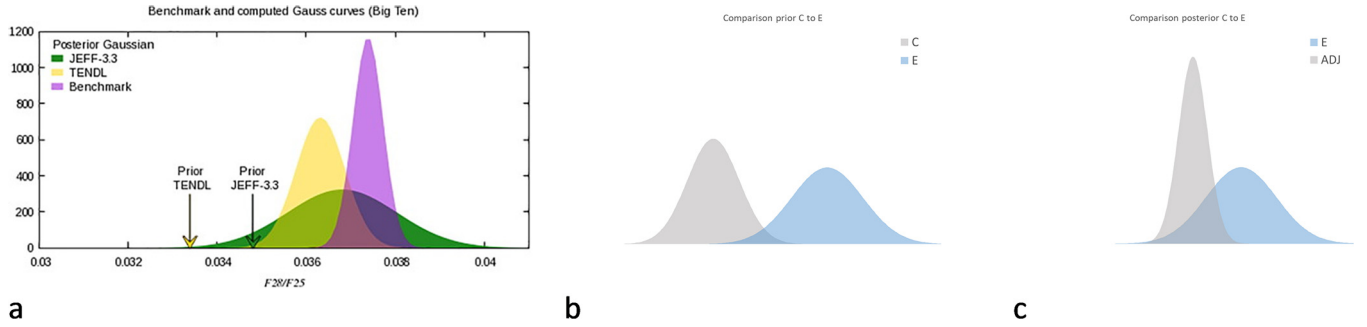


Fig. 1. Qualitative illustration of ND and IEs PDFs [20] (a) and a degree of coincidence between prior (b) and posterior (c) calculational (light grey) and experimental (light-blue) values [7,19].

Summarizing the statements above and bibliography analysis it is easy to see that DA always comprises the following ingredients: (1) objective observations obtained computing representative suitcase of integral experiments data, i.e. calculation-to-experiment ratio for given experiment-based benchmarks, (2) libraries of prior experimental and nuclear data uncertainties needed as the first guess for Data Assimilation process, and (3) relevant Bayesian interference framework that includes, among others, dedicated statistical solvers and parametrized best estimate simulations.

Of course, the DA algorithms in different fields of applications have also different levels of maturity that might be somehow characterized considering the major drawbacks and lessons learned in DA practical implementations [1,13,18].

We are discussing below some examples of good practice and tendencies in DA deployment to characterize in certain extent a technological readiness – maturity – of DA methodologies.

2 Methodological background

As any Kalman filtering, DA has a premise that, given some disagreement between calculated and experimental values, one adjusts parametrized data, performing the best fit of error-weighted expected and observed parameters [1,5,11,17]. It links together such probabilistic categories as conditional probabilities or probability densities $p(\dots | \dots)$, the sets of measurable parameters \mathbf{y} and parameters \mathbf{x} inherent to modeling, and prior information \mathbf{U} from which the prior knowledge on \mathbf{x} is assumed:

$$p(\mathbf{x}|\mathbf{y}, \mathbf{U}) \approx \frac{p(\mathbf{x}|\mathbf{U}) \cdot p(\mathbf{y}|\mathbf{x}, \mathbf{U})}{\int d\mathbf{x} \cdot p(\mathbf{x}|\mathbf{y}) \cdot p(\mathbf{x}|\mathbf{y}, \mathbf{U})}, \quad (1)$$

where the denominator is just a normalization constant [1,17,18]. Then, assuming the high-entropy distributions one builds up the following solving rule to estimate an improved distribution from prior one basing on the

principle of maximum likelihood [1,8,11,18]:

$$\begin{aligned} \text{posterior}[p(\mathbf{x}|\mathbf{y}, \mathbf{U})] &\propto \text{prior}[p(\mathbf{x}|\mathbf{U})] \\ &\times \text{likelihood}[p(\mathbf{y}|\mathbf{x}, \mathbf{U})]. \end{aligned} \quad (2)$$

According to such probabilistic definitions we are considering the results of measurements and calculations as high-entropy distributions represented by relevant³ Probability Density Function (PDF) [12,14]. The coincidence of two values in this case could be graded in terms of an overlapped area (see Fig. 1).

It is easy to see that what DA means is the conversion of prior errors and uncertainties into correction factors and quantified residual uncertainties following the maximum likelihood principle [1,2,8,21,22].

This is why, the kinds of uncertainties we are dealing with – whether they are of a simple error, epistemic or ontological ones [23,24] – dictate which namely technique to be approached to the adjustment, validation or another ill-posed inversed problem.

The first type of uncertainties – *simple error* – appears when prior knowledge would be inconsistent or even wrong. At this level DA clarifies given nuclear reaction models adjusting their parameters by the best fit to pre-selected representative sets of IEs [1,7,8,23].

It requires a robust theoretical model of nuclear reactions⁴ and fully representative IEs that – a few or a large set – should be selected in such a way to discriminate (in a statistical sense) all contributors except for one of interest.

Such demand could be met, inter alia, by design of specific, for example, replacement or oscillation experiments like ones performed at numerous facilities [25], including MINERVE facility [26], IPEN/MB1 [27], and so on.

Talking about the robust theory one could remind, for example, such well-known tools as R-Matrix fitting codes SAMMY [9] or REFIT [10], or CONRAD [11] where the last one could uniformly treat the resolved resonance range (R-Matrix approximation: Reich-Moore and Multi-Level

³ Generally speaking, unknown ones, about which we could have the only some momentums or other aggregated values.

⁴ Like ones based on R-matrix theory or so on.

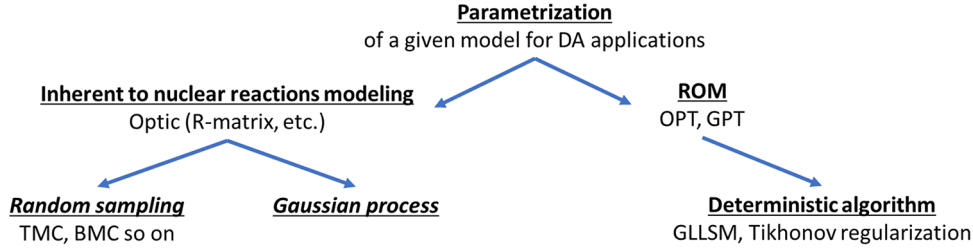


Fig. 2. Parametrization strategy and algorithmic options to be implemented in DA [1–3,5–7,9–12,14,16,18].

Breit-Wigner), the unresolved resonance range (average R-Matrix and Hauser-Feshbach theory) and the fast energy region as it is required by a modern evaluation process.

The second ones — *epistemic uncertainties* — appear due to the imprecise interpolations and extrapolations inherent⁵ to the modeling. Over there, the role of DA would be to not adjust but to provide an evidence-based background⁶ for a science-driven validation.

The last — *ontological* — uncertainties appear due to a different belief system so that the only discovery could resolve such lack of knowledge [19,23]. The role of DA in such case will be the only to contribute in a planning of the further dedicated researches.

2.1 Parametrization strategies

Applying DA, one should take into account that the domain of experiments is discrete and countable while the domain of simulations is continuous and non-countable. Such fundamental discontinuity requires the domain of simulation to be somehow parametrized.

The simplest parametrization strategy is using of ROM to replace physics behind the phenomena of interest by a set of linear response functions, i.e. the following sensitivity coefficients [1,7,28–34]:

$$S_{R,\theta} = \frac{\theta}{R} \cdot \frac{\Delta R}{\Delta \theta} = \frac{\theta}{R} \cdot \left(\frac{\partial R}{\partial \theta} + \sum \frac{\partial R}{\partial \alpha} \cdot \frac{\partial \alpha}{\partial \theta} \right), \quad (3)$$

where $S_{R,\theta}$, R , α , and θ are sensitivity coefficient for a given system's response to a given parameter of nuclear reaction modeling, the system response, nuclear cross section and parameters inherent to nuclear reaction calculations, correspondingly. Two components — explicit $\left(\frac{\partial R}{\partial \theta}\right)$ and implicit $\left(\frac{\partial R}{\partial \alpha} \cdot \frac{\partial \alpha}{\partial \theta}\right)$ — reflect parameters common for different nuclides and reactions parameters [1,2,7,8].

The next kind of strategies presumes the following dependence of the nuclear reactions' model on the countable set variables:

$$\mathbf{Lib}_{ADJ} \approx \mathbf{fun}(\alpha_1, \alpha_2, \dots, \alpha_K | \theta_1, \theta_2, \dots, \theta_L), \quad (4)$$

where \mathbf{Lib}_{ADJ} , $\{\alpha_k : k \in 1 \dots K\}$ and $\{\theta_l : l \in 1 \dots L\}$ are the synthesized/adjusted library, adjusted and non-adjusted parameters, correspondingly.

Apparently, such formalism might give a fully consistent evaluation going deeply to the theoretical models and the best-estimated metamodels inherent to nuclear data calculations. Unfortunately, it would require so many experimental cases making adjustment unaffordable. Nevertheless, the strategy was recognized as sufficient in many practical applications [12,13,15].

Another group represents true (but unknown) library as a weighted superposition of different profiles given as follows:

$$\mathbf{Lib}_{syn} \approx a_1 \cdot \mathbf{Lib}_1 + a_2 \cdot \mathbf{Lib}_2 + \dots + a_N \cdot \mathbf{Lib}_N = \sum_n a_n \cdot \mathbf{Lib}_n, \quad (5)$$

where \mathbf{Lib}_{syn} , \mathbf{Lib}_n and a_n are the desirable synthetic library, n -th generated nuclear data profile (ACE files, typically) and weight factors to be matched, correspondingly, and $n \in 1 \dots N$.

For example, Hierarchic Monte-Carlo [14] operates with such strategy generating the profiles \mathbf{Lib}_n in an iterative randomized process fitting to a given set of IEs data.

Basing on the reasoning above and the bibliography overview we could conventionally consider the variety of DA algorithms as it is shown on Figure 2.

Of course, DA always suffers from the dimensionality of physical model. Indeed, if use discretized modeling in the field of particle transport and reactor physics we are dealing with, as minimum, N_σ parameters determined as a $N_\sigma = N_{IZ} \cdot N_{REA} \cdot N_{EG}$, where N_{IZ} , N_{REA} and N_{EG} are the number of nuclides, the number of channels and the integral number of energy-angular intervals. If use Reich-Moore approximation the dimension will be extended on the numbers of particle and γ -ray channels for each resonance areas.

Since DA assimilation algorithms being different in details all are based on the same theoretical basis we could illustrate the major DA ideas via Master-Equations used in a deterministic methodology like the following one:

$$\frac{\Delta \bar{\sigma}}{\bar{\sigma}} = \widehat{W} \cdot \bar{S}_{IE} \cdot \left(\widehat{V}_{IE} + \widehat{V}_{CLC} + \bar{S}_{IE}^T \cdot \widehat{W} \cdot \bar{S}_{IE} \right)^{-1} \cdot \frac{\Delta \bar{R}_{IE}}{\bar{R}_{IE}}, \quad (6)$$

⁵ Typically, we are covering the scales from femto- to centimeters.

⁶ We are using the term “an evidence-based background” instead of experiments data because in opposite to the values the uncertainty never could be validated against experimental data.

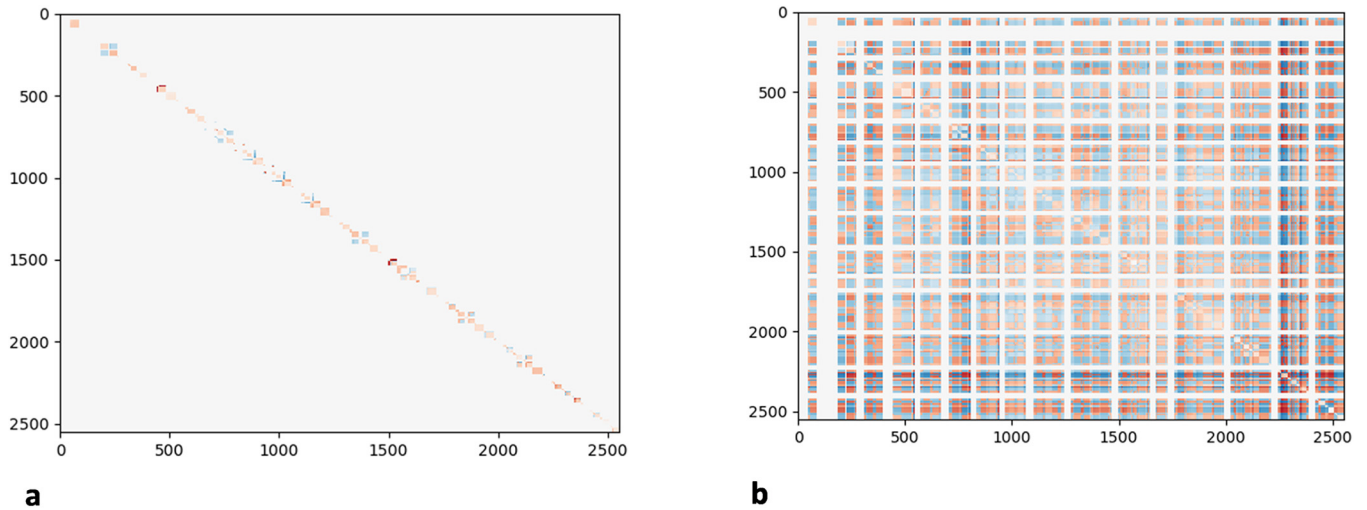


Fig. 3. Prior (a) and posterior (b) covariance matrices for a selected set of nuclide-reactions [19].

where $\left(\frac{\Delta\sigma}{\sigma}\right)$, $\frac{\Delta R_{IE}}{R_{IE}}$, \widehat{W} , \overline{S}_{IE} , \widehat{V}_{IE} and \widehat{V}_{CLC} are a vector of correction factors, a vector of relative discrepancies, prior covariance matrix of nuclear data (CND), calculated vector of sensitivity coefficients for IEs, experimental and calculational covariance matrices, respectively [1,2,10,19].

It gives also a quantified posterior covariance matrix (\widehat{W}') as follows:

$$\widehat{W}' = \widehat{W} - \left(\overline{S}_{IE}^T \cdot \widehat{W} \cdot \overline{S}_{IE}\right) \cdot \left(\widehat{V}_{IE} + \widehat{V}_{CLC} + \overline{S}_{IE}^T \cdot \widehat{W} \cdot \overline{S}_{IE}\right)^{-1} \times \left(\overline{S}_{IE}^T \cdot \widehat{W} \cdot \overline{S}_{IE}\right), \quad (7)$$

where all notations are similar to given above [1,2,10,19].

One can see that posterior covariance matrix (\widehat{W}') and, therefore, posterior uncertainties for a Quantity of Interest (QoI) do not depend on calculation-to-measurement discrepancies; while correction factors $\left(\frac{\Delta\sigma}{\sigma}\right)$ do and the only vector of sensitivities (\overline{S}_{IE}) represents the physics behind the IEs and the applications.

Available nowadays continues energy and arbitrary geometry Monte-Carlo sensitivity analysis allows performing fine-resolution adjustment as it is demonstrated in such tools as SAMINT (nuclear data adjustment with SAMMY based on Integral Experiments) [9,16] which complements Bayesian fit performed using SAMMY (multilevel R-matrix fits to neutron and charged-particle cross-section data using Bayes' equations) tool.

2.2 Integral experiments data and an evidence-based background

As said, the adjustment critically depends on a quality of IEs data, including consistency of their uncertainties and covariance (see component \widehat{V}_{IE} in Eqs. (6) and (7)).

These uncertainties and experimental covariance matrices are resulted from the physics-based evaluation

of measurements, as such, and the experimental conditions in the manner similar to what has been implemented in the International Criticality Safety Benchmark Experiments Project (ICSBEP) and the International Reactor Physics Experiments (IRPhE) Project, for example [7,25].

Historically, IEs were considered, mainly, as mock-ups allowing to study major characteristics of nuclear systems, optimizing and examining reactor control systems, radiation shielding and others using zero or low power facilities to minimize all risks associated with nuclear safety and radiation protection. Nowadays, due to a progress in numerical simulations and increased requirements to an accuracy of modeling, such vision, except for very rare cases, seems to be obsolete.

Of course, experiments would be of different kinds, including criticality and reactivity studies, reaction rates measurements, depletion analysis and so on. The only required experimental data to be stringently evaluated. Unfortunately, we have to note that the experimental covariances are scarcely available even in the popular Handbooks.

2.3 Information content of the posterior bias and uncertainties

The next essential ingredient of DA – prior uncertainties – is crucial for any Bayesian inference technique [1,19,35–37].

Historically, nuclear data uncertainties are available in several group-wise matrix formats associated with the most popular libraries such as JEFF, JENDL, ENDF, TENDL, SCALE, etc. In this context we could mention such covariances data libraries as BOLNA created in a collaboration among BNL, ORNL, LANL, NRG, ANL, COMMARA-2.0 derived for ENDF-based nuclear data in one of OECD-NEA project and COMACV1 and others attached to JENDL, TENDL and SCALE projects [35,38].

In the past the covariance matrices have been based, largely, on the expert judgements. Albeit today extensive world-wide efforts were mounted to determine the scientific basis to establish relevant CND (Fig. 3).

However, posterior CND – generated after adjustment – never fully inherits the prior CND. In fact, DA integrate somehow an information brought with the used IEs onto corrected nuclear parameters and their uncertainties [35]. In the years of DA practical implementations an intensive discussion arisen on the question how to interpret an appearance in posterior CND cross-covariance members have not been in prior ones [36]. It was found that the cross-covariance members in a posterior CND always contain the traces of IEs data [1,19,36,37], characterizing in certain extent an efficiency of the adjustment [19].

3 Best practice in data assimilation worldwide

From the very beginning, nuclear technological science intended all concepts and statements to have a solid basis in reality. In all domains of nuclear engineering from design to safety regulation, it seems crucial having access to objective observations, including operational background, basic and dedicated Integral Experimental (IE) programs [39–44].

However, we could use both legacy and newly established IEs to improve or to validated nuclear data libraries. The only issue is that we have to unfold somehow the IEs data using them in a nuclear data evaluation process. It is possible to do if IEs are numerous, their set is statistically significant and there is a robust DA approach consistent for a given field of interest [1,2,8,19].

From more general points of view, one might distinguish three the following major groups of DA practical applications: (1) simple data adjustment contributing to problem-oriented and general-purpose libraries [1,2,4,5,7], (2) science-driven validation of nuclear data libraries and simulations [2,6,19], and (3) knowledge-based prioritization of dedicated basic research programs [19].

As far as DA techniques have different backgrounds for different applications one seems reasonable to characterize them below in terms of the level of maturity.

3.1 1st application: simple data adjustment

As said, the very basic idea of nuclear data adjustment is an optimized fit of the modeling of nuclear reaction or of nuclear systems to the well-evaluated consistent and credible IEs data [1]. Such simple adjustment requires to use fully representative [2,39] sets of high-fidelity IEs data [25].

In the past, in the early 70s, one of such criteria of representativity factor ($r_{IE,QoI}$) was derived as follows:

$$r_{IE,QoI} = \frac{(\bar{S}_{QoI}^T \cdot \widehat{W} \cdot \bar{S}_{IE})}{\sqrt{(\bar{S}_{QoI}^T \cdot \widehat{W} \cdot \bar{S}_{QoI}) \cdot (\bar{S}_{IE}^T \cdot \widehat{W} \cdot \bar{S}_{IE})}}, \quad (8)$$

where experiments assumed to be independent and all notations are similar to ones given above [1,39]. In case of correlated experiments, one should implement one or

another iterative process to quantify a representative factor for each single experiment [19].

Furthermore, historically, the DA in nuclear engineering have been applied in two the following axes: (1) to generate data libraries adjusted to a given set of applications like, for example, ERALIB1 [4] and early versions of ABBN library [1], and (2) to refine knowledge on certain parameters of nuclear reaction models [5,9,10,12].

Practically, there are only two major ideas of an adjustment: (1) to fit some aggregated parameters like group-wise cross sections and, then, to refine the adjusted integral values, correcting the very basic parameters of the nuclear process model; (2) to correct these parameters directly fitting the models of nuclear processes to IEs data.

One can see that in both axes Data Assimilation demonstrated maturity sufficient for current requirements of nuclear data evaluations [7,38,40,45].

3.2 2nd application: science-driven V & UQ

Together with corrected ND section DA quantifies their uncertainties generating posterior CNDs. It could be used evaluating the quality of the adjustment process as well as validating the nuclear data libraries. Such application – to support a validation process – may become even more important than the data correction. Indeed, we have a few well-elaborated and recognized brands of nuclear data projects (ENDF/B, JENDL, JEFF, BROND, ROSFOND, CENDL and TENDL, and some others [7,38]). It seems unlikely to repeat or improve any of them by a single design or scientific organization but they could be characterized in terms of anticipated uncertainties in the field of users' interests.

The Validation through Uncertainty Quantification requires DA algorithms to be, mainly, robust and the only on the second order to be of high resolution.

It should be noted that a science-driven validation – that is exactly our case – separates domains of validation and applications. It means that we can use whatever kinds of experiments – critical, reactivity, reaction rates and so on – to estimate biases and uncertainties for any Quantity of Interest (QoI). What is needed is to have relevant sensitivity coefficients or functional models to be combined with corrections and posterior CNDs. Thus, in terms of GLLS methodology the bias of QoI could be computed as follows:

$$\frac{\Delta QoI}{QoI} = \bar{S}_{QoI}^T \cdot \frac{\Delta \bar{\sigma}'}{\bar{\sigma}'}, \quad (9)$$

where $\frac{\Delta QoI}{QoI}$ and \bar{S}_{AO} area scalar bias predicted by DA and a vector of sensitivity coefficient for the QoI, correspondingly; and the uncertainties as

$$\delta_{QoI}^2 = (\bar{S}_{AO}^T \cdot \widehat{W}' \cdot \bar{S}_{AO}), \quad (10)$$

where δ_{QoI} is the relative standard deviation when other notations are similar to given above.

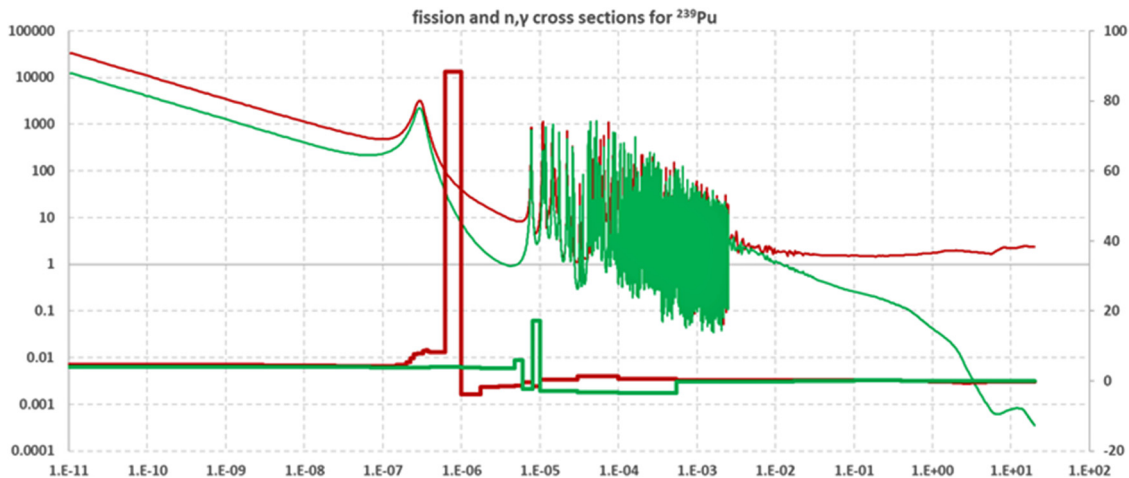


Fig. 4. Energy spanned cross sections and relative data gain indicators (XS_{ADDED}).

3.3 3rd application: step toward ontological uncertainty treatment

While the first two groups of DA were well-illustrated by practical cases, conference and journal papers [1,7] the third group – what to do if we are dealing with an ontological issues – was not pronounced so far. As said, neither adjustment or comparison with observations but the only a kind of discovery would help treating an ontological uncertainty. However, even in this case DA could become useful bounding the impact of such kind of uncertainties and contributing in an establishment of further problem-oriented basic research programs [19].

For example, years ago, the nuclear criticality safety community considered one hypothetical case of a criticality on a fuel powder-mixing apparatus as one of high priority. Physically, the configuration to be assessed was a moisture in the mixture of reactor-grade plutonium and uranium oxides and the critical conditions were reached with an epithermal spectrum. Because of many reasons, the number of representative integral experiments was very limited while a few available give a discrepancy on several percent of k_{eff} , that correspond to one-third or, even, one-half of critical mass. Later, the dedicated parametric experimental program was established with under-moderated ^{240}Pu containing critical assemblies [19,25,41,42]. As a result, the experiments confirmed an existence of the issue while used, then, DA helped characterizing specific safety margins by posterior bias and uncertainties [19] remaining, though, unclear which namely nuclide-reaction led to these discrepancies.

By chance, in this particular case, we have had two sets of IEs data. The first one – the “basic set” of experiment-based benchmarks – taken from ones available in the Handbooks [25]. The second – complemented – set of the same “basic” ones complemented by the newly obtained ones. Comparing two correction factors derived from these two sets we could estimate the following vector of indicators XS_{ADDED} to point down the nuclide, reaction

and energy interval “responsible” for such discrepancy:

$$XS_{ADDED} = 1 - \frac{[\frac{\Delta\sigma}{\sigma}]_{ADDED}}{[\frac{\Delta\sigma}{\sigma}]_{BASIC}}, \quad (11)$$

where $[\frac{\Delta\sigma}{\sigma}]_{BASIC}$ and $[\frac{\Delta\sigma}{\sigma}]_{ADDED}$ are the factors adjusted using basic and complementary sets of benchmarks, correspondingly. In our case the energy spanned XS_{ADDED} profile depicted on Figure 4 shows that the field to be elaborated, most probably, relates to a right wing of 0.296 eV fission resonance on ^{239}Pu .

It should be noted, that this conclusion has been surprisingly confirmed by an interpretation of some recent tests of modern ND libraries against a fuel depletion experimental benchmark associating the issue with a right wing of the first ^{239}Pu fission resonance⁷.

Using DA [19] we revealed the questionable area to lie below eV in a total contradiction with the intuitive statements that this area to be validated using experiments with thermal spectra [42,43].

4 Discussion: technology readiness level

Assessing the maturity of DA algorithms, we divided them conventionally by three groups like: (1) ROM/ROM, (2) linear-precise and (3) precise-precise representations [1–3,5,7–9,12,14]. The analysis is presented on Figure 5 by applications – ND adjustment, ND Validation through Uncertainty Quantification and contribution in basic research planning – and by the groups of algorithms. The bigger relevant circle on the Figure the higher level of maturity.

The first axe – ROM/ROM – means that the models of nuclear reactions and particle transport simulations were replaced by their Reduced Order Modeling analogous such as relevant sensitivity coefficients. The nuclear reaction

⁷ Discussion within one of OECD-NEA/NDB projects and personal communications with Luiz Leal (IRSN).

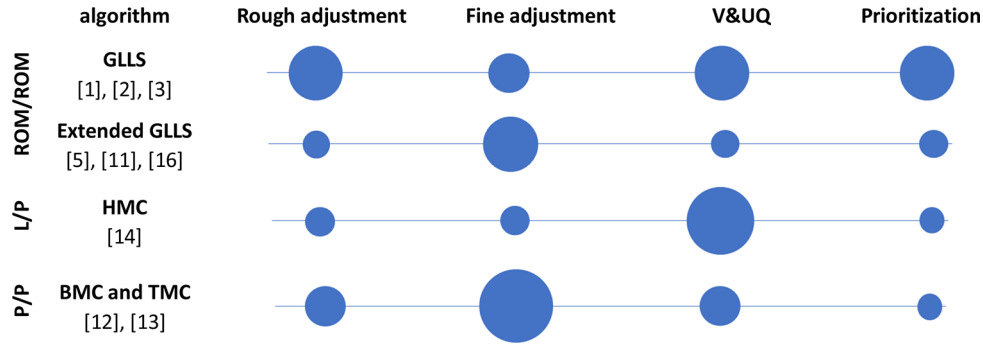


Fig. 5. Maturity (Technology Readiness Level) practically demonstrated using DA algorithms with respect to ND adjustment, validation and relevant studies planning.

model (first abbreviation) was represented as a set of group-wise cross sections [1–3] including, normally, micro-data with Wescott g -factors and Bondarenko f -factors and, if possible, vectors of sub-groups or by the parameters inherent to the high-fidelity nuclear reaction modeling [5,9,11,18]. The particle transport model (second abbreviation) was also given as a set of group-wise sensitivity coefficients, comprised, of course, explicit and implicit components of sensitivity [1,2,7]. In ROM/ROM biases and uncertainties can be used immediately [6,19], while ND correction factors should be somehow unfolded and assessed [1,16]. Thus, we believe that the DA maturity over here seems to be sufficient as for data adjustment as for validation.

The second axe—precise/precise (P/P)—means high-fidelity or, even, precise modeling as for reactor physics as for nuclear reactions. Apparently, it could generate fully balanced and adjusted libraries. However, it still seems unclear how to adjust some pre-calibrated semi-empiric elements contained in high-fidelity theoretical model intended to nuclear reactions calculations.

The third axe — linear/precise (L/P) —represents nuclear data as a superposition of pre-generated high-fidelity profiles. It is usually associated with a Hierarchical Monte-Carlo being oriented, mainly, on validation [14].

In addition, we identified some bottlenecks for DA. First of all, in terms of DA methodologies, one still needs to elaborate an adjustment for composed — non-linear — operators like, for example, one of fission production where ν -bar and PNFS are correlated ($\frac{1}{4\pi} \cdot \chi \cdot \nu \cdot \Sigma_{fiss} \leftarrow \dots \chi(\nu) \cdot \nu \dots$).

Concerning the IEs, one still needs in high-fidelity experimental covariance matrices that exist but not numerous enough in the Handbooks and do not exist for different functionals, like covariance between the measurements of reaction rates and kinetic parameters, for example.

Finally, one should note that some IEs were applied in the ND tuning. These experiments have to be withdrawn from the adjustment and validation or, at least, users should be informed about them.

5 Conclusions

Data Assimilation belongs, mainly, to a field of information technology, is presented in the nuclear technological

science from the sixties of the last century. Known as Nuclear Data adjustment, it was providing users with so-called design-oriented multi-groups libraries and so on.

Among others, the use of the adjustment was warranted if nuclides to be studied were rare or short-lived, or dangerous complicating or, even, making impossible any differential measurements.

Nowadays, despite or, may be, due to a notable success of Data Assimilation there are no more room for rough adjustment, because of enhanced requirements to Nuclear Data accuracy. We are talking either about the fine Nuclear Data calibration via parameters inherent to the nuclear reaction modeling or about the science-driven Validation through Uncertainty Quantification where we could use any, even rough, Data Assimilation algorithm.

As said, Data Assimilation contributes to a nuclear data evaluation combining differential and integral experiments data. In this case we are dealing with the simple errors — discrepancies between calculated and experimental values — and their covariances in order to generate optimally balanced problem-oriented libraries.

Eventually, Data Assimilation could substantiate a science-driven Validation providing assessor with an evidence-based background.

In addition, Data Assimilation could be used in gap analysis somehow contributing in an establishment of dedicated basic research programs.

Further development of Data Assimilation for Nuclear Data evaluation could be considered, among others, by the following axes: (1) an extension of the applications enhancing comprehensive optimization and validation of nuclear data libraries; (2) an improvement of numerical algorithms involving recently developed data science techniques; and (3) an elaboration of experimental databases.

Summarizing the reasoning above we could conclude that Data Assimilation, as an approach, has a sufficient maturity for the nuclear engineering applications having, at the same time, significant potential for further refinement.

The paper is written in a memory of Dr. Massimo Salvatores. He, among others, made a great contribution to the rise of a Perturbation theory and Data Assimilation involvement in a

broad range of scientific domains of nuclear engineering, including reactor physics and control, innovative technologies and, on a top of this, in nuclear data evaluation and validation.

We would also extend our appreciation to OECD/NEA staff and expert groups' members for their deep involvement in a scientific discussion on a role and practical implementation of Bayesian-based methodologies in a nuclear technological science.

Author contribution statement

- 1) Evgeny Ivanov: general coordination and contribution in all the chapters.
- 2) Cyrille De Saint-Jean: contribution in the chapters of ethodological background, discussion and conclusions, and in the bibliography.
- 3) Vladimir Sobes: principle contribution in the chapters of methodological background, best practice in Data Assimilation, discussion and conclusions.

References

1. M. Salvatores et al., Methods and issues for the combined use of integral experiments and covariance data: results of a NEA international collaborative study, Nucl. Data Sheets **118**, 38–71 (2014)
2. L.N. Usachev, Y. Bobkov, Planning on optimum set of microscopic experiments and evaluations to obtain a given accuracy in reactor parameter calculations, INDC CCP-19U, IAEA International Nuclear Data Committee (1972)
3. J.L. Rowlands, L.D. Macdougall, The use of integral measurements to adjust cross-sections and predicted reactor properties, Proceedings of the International Conference on Fast Critical Experiments and their Analysis, ANL-7320 (1966)
4. E. Fort, G. Rimpault, J-C. Bosq et al., Improved performances of the fast reactor calculational system ERANOS-ERALIB1 due to improved a priori nuclear data and consideration of additional specific integral data, Ann. Nucl. Energy **30**, 1879–1898 (2003)
5. C. de Saint-Jean (Co-ordinator), Assessment of Existing Nuclear Data Adjustment Methodologies, Report by the Working Party on International Evaluation Co-operation of the NEA Nuclear Science Committee, Vol. 33, NEA/WPEC-33, OECD/NEA, 2011
6. G. Palmiotti, M. Salvatores, The role of experiments and of sensitivity analysis in simulation validation strategies with emphasis on reactor physics, Ann. Nucl. Energy **52**, 10–21 (2013)
7. Dan Gabriel Cacuci (ed.) Handbook of nuclear engineering: Vol. 1: nuclear engineering fundamentals, Springer, Boston, MA (Springer 2010)
8. G. Evensen, Data Assimilation: The Ensemble Kalman Filter (Springer, Berlin, 2006)
9. N.M. Larson, ORNL Report ORNL/TM-9179/R8, 2008
10. M. Moxon et al., UKNSF Report, 2010
11. P. Archier, C. De Saint Jean, O. Litaize, G. Noguère, L. Berge, E. Privas, P. Tamagno, CONRAD evaluation code: development status and perspectives, Nucl. Data Sheets **118**, 488–490 (2014)
12. A.J. Koning, Bayesian Monte Carlo method for nuclear data evaluation, Nucl. Data Sheets **123**, 207–213 (2015)
13. D. Siefman, M. Hursin, D. Rochman et al., Stochastic vs. sensitivity-based integral parameter and nuclear data adjustments, Eur. Phys. J. Plus **133**, 429 (2018)
14. A. Hoefler et al., MOCABA: A general Monte Carlo–Bayes procedure for improved predictions of integral functions of nuclear data, Ann. Nucl. Energy **77**, 514–521 (2015)
15. C. De Saint Jean, P. Archier, E. Privas, G. Noguere, On the use of Bayesian Monte-Carlo in evaluation of nuclear data, EPJ Web Conf. **146**, 02007 (2017)
16. V. Sobes, L.C. Leal, G. Arbanas, Nuclear data adjustment with SAMMY based on integral experiments, Anaheim, California **111**, 843–845 (2014)
17. C. de Saint Jean, P. Archier, E. Privas, G. Noguère, O. Litaize, P. Leconte, Evaluation of cross section uncertainties using physical constraints: focus on integral experiments, Nucl. Data Sheets **123**, 178 (2015)
18. W. Martienssen (ed.), Low Energy Neutron Physics, Landolt-Bornstein (Springer-Verlag, Berlin, 2000)
19. T. Ivanova, E. Ivanov, I. Hill, Methodology and issues of integral experiments selection for nuclear data validation, EPJ Web Conf. **146** (2017)
20. S. Pelloni, D. Rochman, Performance assessment of adjusted nuclear data along with their covariances on the basis of fast reactor experiments, Ann. Nucl. Energy **121**, 361–373 (2018)
21. M.G. Kendall, A. Stuart, in The Advanced Theory of Statistics. Vol. 2, Inference and Relationship (Hafner, New York, 1961), pp. 474–483
22. V.F. Turchin et al., The use of mathematical-statistics methods in the solution of incorrectly posed problems, Soviet Physics Uspekhi **13**, 681 (1971)
23. K. Beven, Facets of uncertainty: epistemic uncertainty, nonstationarity, likelihood, hypothesis testing, and communication, Hydrol. Sci. J. **61**, 1652–1665 (2016)
24. W. Oberkampf, C. Roy, Verification and Validation in Scientific Computing (Cambridge University Press, UK, 2010)
25. J.B. Briggs, J.D. Bess, J. Gulliford, Integral benchmark data for nuclear data testing through the ICSBEP & IRPhEP, Nucl. Data Sheets **118** (2014)
26. A. Santamarina et al., Reactivity worth measurement of major fission products in MINERVE LWR lattice experiment, Nucl. Sci. Eng. **178**, 562–581 (2014)
27. L. Leal, A.D. Santos, E. Ivanov, T. Ivanova, Impact of ^{235}U resonance parameter evaluation in the reactivity prediction, Nucl. Sci. Eng. **187**, 127–141 (2017)
28. K.F. Raskach, An Improvement of the Monte Carlo generalized differential operator method by taking into account first- and second-order perturbations of fission source, Nucl. Sci. Eng. **162**, 158–166 (2009)
29. B.C. Kiedrowski, F.B. Brown, P.P.H. Wilson, Adjoint-weighted tallies for k-eigenvalue calculations with continuous-energy Monte Carlo, Nucl. Sci. Eng. **168**, 226–241 (2011)
30. M. Aufiero et al., A collision history-based approach to sensitivity/perturbation calculations in the continuous energy Monte Carlo code SERPENT, Ann. Nucl. Energy, Volume **85**, 245–258 (2015)

31. E. Brun et al., TRIPOLI-4[®], CEA, EDF and AREVA reference Monte Carlo code, *Ann. Nucl. Energy* **82**, 151–160 (2015)
32. A. Jinaphanh, N. Leclaire, Continuous-energy perturbation methods in the MORET 5 code, *Ann. Nucl. Energy* **114**, 395–406 (2018)
33. C.M. Perfetti, B.T. Rearden, W.R. Martin, SCALE continuous-energy eigenvalue sensitivity coefficient calculations, *Nucl. Sci. Eng.* **182**, 332–353 (2016)
34. M.T. Pigni, M. Herman, P. Oblozinsky, F.S. Dietrich, Sensitivity analysis of neutron total and absorption cross sections within the optical model, *Phys. Rev.* **C83**, 24601 (2011)
35. I. Kodeli, Comments on the status of modern covariance data based on different fission and fusion reactor studies, *EPJ Nuclear Sci. Technol.* **4**, 46 (2018)
36. G. Palmiotti, M. Salvatores, Cross section covariances: a user perspective, *EPJ Nuclear Sci. Technol.* **4**, 40 (2018)
37. E. Bauge, S. Hilaire, P. Dossantos-Uzarralde, Evaluation of the covariance matrix of neutronic cross sections with the Backward-Forward Monte Carlo method, *Inter. Conf. Nucl. Sci. Technol.* 259–264 (2007)
38. M.B. Chadwick et al., CIELO collaboration summary results: international evaluations of neutron reactions on uranium, plutonium, iron, oxygen and hydrogen, *Nucl. Data Sheets* **148**, 189–213 (2018)
39. D. Kumar, S.B. Alam, H. Sjöstrand, J.M. Palauand, C. De Saint Jean, Influence of nuclear data parameters on integral experiment assimilation using Cook’s distance, *EPJ Web Conf.* **211**, 07001 (2019)
40. C. De Saint Jean et al., Evaluation of neutron-induced cross sections and their related covariances with physical constraints, *Nucl. Data Sheets* **148**, 383–419 (2018)
41. T.T. Ivanova, M.N. Nikolaev, K.F. Raskach, E.V. Rozhikhin, A.M. Tsiboulia, Use of international criticality safety benchmark evaluation project data for validation of the ABBN cross-section library with the MMK-KENO code, *Nucl. Sci. Eng.* **145**, 247–255 (2003)
42. The Need for Integral Critical Experiments with Low-moderated MOX Fuels, in *Proceedings of the Workshop, Paris, France, 14–15 April 2004*, OECD NEA, No. 5668, ISBN 92-64-02078-0
43. L.C. Leal, G. Noguere, C. de Saint Jean, A.C. Kahler, ²³⁹Pu resonance evaluation for thermal benchmark system calculations, *Nucl. Data Sheets* **118**, 122–125 (2014)
44. J.B. Clarity, W.J. Marshall, B.T. Rearden, I. Duhamel, Selected uses of TSUNAMI in critical experiment design and analysis. In: *2020 ANS Virtual Winter Meeting, Transactions, Volume 123, Number 1, 2020*, pp. 804–807
45. T. Frosio, T. Bonaccorsi, P. Blaise, Extension of Bayesian inference for multi-experimental and coupled problem in neutronics – a revisit of the theoretical approach, *EPJ Nuclear Sci. Technol.* **4**, 19 (2018)

Cite this article as: Evgeny Ivanov, Cyrille De Saint-Jean, Vladimir Sobes, Nuclear data assimilation, scientific basis and current status, *EPJ Nuclear Sci. Technol.* **7**, 9 (2021)