

Influence of nuclear data parameters on integral experiment assimilation using Cook's distance

D. Kumar^{1,2}, S. B. Alam¹, H. Sjöstrand², J. M. Palau¹, and C. De Saint Jean^{1,*}

¹CEA, DEN, Cadarache, F-13108 Saint Paul les Durance, France

²Department of Physics and Astronomy, Uppsala University, Sweden

Abstract. Nuclear data used in designing of various nuclear applications (*e.g.*, core design of reactors) is improved by using integral experiments. To utilize the past critical experimental data to the reactor design work, a typical procedure for the nuclear data adjustment is based on the Bayesian theory (least-square technique or Monte-Carlo). In this method, the nuclear data parameters are optimized by the inclusion of the experimental information using a Bayesian inference. The selection of integral experiments is based on the availability of well-documented specifications and experimental data. Data points with large uncertainties or large residuals (outliers) may affect the accuracy of the adjustment. Hence, in the adjustment process, it is very important to study the influence of experiments as well as of the prior nuclear data on the adjusted results. In this work, the influence of each individual reaction (related to nuclear data) is analyzed using the concept of Cook's distance. First, JEZEBEL (Pu239, Pu240 and Pu241) integral experiment is considered for data assimilation and then the transposition of results on ASTRID fast reactor concept is discussed.

1 Introduction

Nuclear data contains all the nuclear physics needed for the mathematical modeling of a nuclear application (*e.g.*, for reactor design). Nuclear data is evaluated using nuclear reaction models in combination with differential data and is improved by using integral experiments [1, 2]. The improvement using integral experiments is called nuclear data adjustment which is typically based on the Bayesian theory, where the nuclear data parameters are optimized by the inclusion of experimental information using least square analysis or Monte-Carlo approach [2]. Parameters in the context of nuclear data can refer to both, the underlying nuclear model parameters [3], or to the nuclear data itself. In the context of this paper, nuclear data parameters only refer to the multi-group values of the nuclear data defined by isotope, energy and nuclear observable (cross-section of a certain reaction, $\bar{\nu}$, etc).

The so-called “adjusted data” has been used in the design of fast reactors such as Phenix and Super-Phenix in France [4, 5]. However, the adjustment process has raised several questions about the physical meaning of the individual adjustment (*i.e.* by isotopes, reaction types and energy ranges). There is no clear definition of the application domain of the adjusted multi-group data. For a new reactor concept, it is difficult to say if a previously adjusted data can be used or a new adjustment will be needed with a different set of integral experiments.

*Corresponding author: e-mail: cyrille.de-saint-jean@cea.fr

How to choose a set of integral experiments for a particular reactor concept is also not obvious. A detailed analysis is needed to estimate the influence of input data on the adjustment. These data points could be a given observation (integral experiment) or a given nuclear data parameter. The Cook's distance is used to estimate the influence of a data point in performing the least square fit [6, 7]. In this work, the influence of input parameters on the Bayesian fit is analysed using the concept of Cook's distance.

2 Nuclear data adjustment using Bayesian inference

The Bayesian inference is a mathematical framework that allows one to assimilate information from measurements (microscopic and integral) and from a prior knowledge of the parameters, in order to reduce the uncertainties of these parameters and modify their values if necessary.

2.1 General principle

The Bayes' theorem gives the relationship between the probability densities of the posterior, prior knowledge and measurements. If $\vec{y} = y_i (i = 1, 2, \dots, N_y)$ are some experimentally measured variables and \vec{x} are the parameters defining the model M , which is used to simulate these variables, the probability of the adjusted parameters can be estimated using the Bayesian inference. Using the Bayes' theorem, one can write the following relationship between the posterior, prior and the likelihood:

$$p(\vec{x}|\vec{y}, U) = \frac{p(\vec{x}|M, U).p(\vec{y}|\vec{x}, M, U)}{\int d\vec{x}.p(\vec{x}|M, U).p(\vec{y}|\vec{x}, M, U)} \quad (1)$$

where the probability density of these parameters $p(\vec{x}|M, U)$ is called a prior of the background (the prior knowledge U). $p(\vec{y}|\vec{x}, M, U)$ is the likelihood of the measurements knowing \vec{x} and $p(\vec{x}|\vec{y}, U)$ is the posterior probability density of the parameters. The denominator term in the above equation is a normalization constant.

This relation can therefore be understood as the advancing of the a prior knowledge by a likelihood function of new measurements, *i.e.*,

$$posterior [p(\vec{x}|\vec{y}, U)] \propto prior [p(\vec{x}|U)] .likelihood [p(\vec{y}|\vec{x}, U)] \quad (2)$$

A fitting procedure can be used (*e.g.*, generalized least square fit) to estimate the first two moments of the posterior or to obtain the posterior distribution of these parameters. In nuclear data field, this is a standard procedure for nuclear data adjustment [2]. In the context of this paper, x is the multi-group nuclear data, M is a reactor model and y is the result of an integral experiment.

2.2 Analytical resolution

The above problem can be solved analytically by considering some necessary assumptions. In particular, the prior probability density and the likelihood are considered as Gaussian multi-parameterized distributions. Thus, if the prior data and its covariance matrix are respectively \vec{x}_m and M_x , and \vec{y} is a set of integral experiments with M_y the covariance matrix, the posterior density of \vec{x} can be written as:

$$\text{posterior} [p(\vec{x}|\vec{y}, U)] \propto e^{-(1/2)((\vec{x}-\vec{x}_m)^T M_x^{-1}(\vec{x}-\vec{x}_m) + (\vec{y}-\vec{t})^T M_y^{-1}(\vec{y}-\vec{t}))} \quad (3)$$

where \vec{t} is the theoretical model prediction (or result from computer simulation), which is functions of \vec{x} . For posterior distribution, the Laplace approximation is used for the deterministic resolution of the least-square equation which consists in considering that this density is also of the Gaussian type is used.

Now, if the mean and the covariance matrix of the prior parameters, and the experimental covariance matrix are provided, the parameters after adjustment (and the posterior covariance matrix) can be estimated by looking for the minimum the following cost function:

$$\chi_{GLS}^2 = (\vec{x} - \vec{x}_m)^T M_x^{-1} (\vec{x} - \vec{x}_m) + (\vec{y} - \vec{t})^T M_y^{-1} (\vec{y} - \vec{t}) \quad (4)$$

Using the least square fitting technique, the solution of this minimization gives the new adjusted values :

$$\vec{x} - \vec{x}_m = M_x G^{(n)T} (M_y + G^{(n)} M_x G^{(n)T})^{-1} \cdot (\vec{y} - t(\vec{x}_m)) \quad (5)$$

and the a posterior covariance matrix:

$$\tilde{M}_x = [M_x^{-1} + G^{(n)T} M_y^{-1} G^{(n)}]^{-1} \quad (6)$$

where G is the sensitivity matrix, *i.e.*, a matrix of derivatives. n is associated with the n^{th} iteration of the algorithm corresponding to a given numerical convergence. The entries of the sensitivity matrix are calculated as:

$$(G)_{ij} = \left(\frac{\partial t_i}{\partial x_j} \right) \quad (7)$$

with t_i the theoretical calculation corresponding to the experimental value. The coefficients G_{ij} are computed using sensitivity analysis (perturbation theory). For the adjustment of multi-group nuclear data, the model parameters x and the covariance matrix M_x in the above equations can be replaced by the nuclear data (σ_p) and its covariance matrix (M_{σ_p}) respectively.

3 Effect of influential data using Cook's distance

It is commonly accepted that the effect of input data on regression models should be analysed. A strongly influential case dominates the regression model in such a way, that the estimated regression line lies close to this case. The analysis of residuals cannot be used to detect influential cases as the cases with high residuals are indicated as outliers. The influence of an individual data on the overall regression fit can be thought of in terms of how much the fit would differ if the individual data is not included in the fitting process. If the predictions are the same with or without the individual data, the latter has no influence on the regression model. If the predictions differ significantly when the observation is not included in the analysis, the data is highly relevant.

Cook's distance provides an overall measurement of the change in all parameter estimates. The concept of Cook's distance is a good measure of the influence of observation and is proportional to the sum of the squared differences between predictions made with all observations in the analysis and predictions made by removing the observation in question. A formula for the Cook's Distance is provided in [6, 7]. In this work, the effect of the inputs is analysed using the concept of Cook's distance.

If σ_p is the original posterior nuclear data and σ_p^i is the adjusted posterior nuclear data when the i^{th} parameter is discarded in the adjustment process, the Cook's distance can be calculated as (see [7]):

$$D_i = (\sigma_p^i - \sigma_p) M_{\sigma_p}^{-1} (\sigma_p^i - \sigma_p) \quad (8)$$

where M_{σ_p} is the covariance matrix of the posterior parameters. The M_{σ_p} matrix is not re-estimated when a subset of data is removed in calculating the Cook's distance. The covariance matrix is calculated using uncertainties (the standard deviation) of the nuclear data parameters and their correlation matrix. However, due to the high dimensional nuclear data used in the adjustment process, currently, several issues (such as singularity, negative definite) are encountered when inverting the covariance matrix. For those reasons, currently, a modified covariance matrix based on the nuclear data uncertainties is considered in calculating the Cook's distance for this work.

4 Test case

For this study, the prior nuclear data uncertainties and their correlations are taken from CEA files COMAC-V1 in 33 energy groups [2]. The JEZEBEL experiment is considered as an integral experiment for the adjustment [8]. The JEZEBEL experiment was a small spherical assembly of plutonium alloyed with gallium (3.41% Ga, 91.95% Pu239, 4.35% Pu240 and 0.29% Pu241). Pu239 was the major component of the core in this experiment. It was used to determine the critical mass of spherical and homogeneous Pu-alloy. Using Cook's distance, the influence of isotopes and nuclear reactions on the adjustment is investigated in this section.

4.1 Influence of Pu isotopes on the adjustment

To analyse the effect of Pu isotopes on the adjustment, Pu239, Pu240 and Pu241 isotopes are considered in the adjustment process. It means that the JEZEBEL (integral experiment) sensitivities of k_{eff} with respect to all Pu239, Pu240 and Pu241 nuclear reactions are provided in the adjustment process. The Cook's distance is plotted in Figure 1. The Cook's distance for Pu239 is very large in comparison to Pu240 and Pu241 isotopes. It indicates that the major contribution on the adjustment is due to the k_{eff} sensitivities of the Pu239 isotope, which is justified from the core composition of the JEZEBEL experiment. It is also important to investigate the effect of individual nuclear reaction cross-section of Pu239 isotope on the adjustment, which is analysed in the next section.

4.2 Influence of Pu239 nuclear reactions on the adjustment

To analyse the effect of Pu239 nuclear reactions on the adjustment, the JEZEBEL sensitivities of k_{eff} with respect to all Pu239 nuclear reactions are provided in the adjustment process.

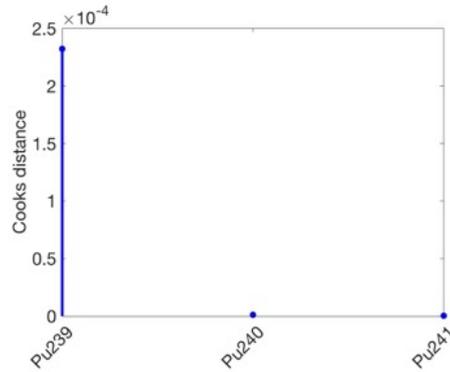


Figure 1. Cook's distance for Pu isotopes

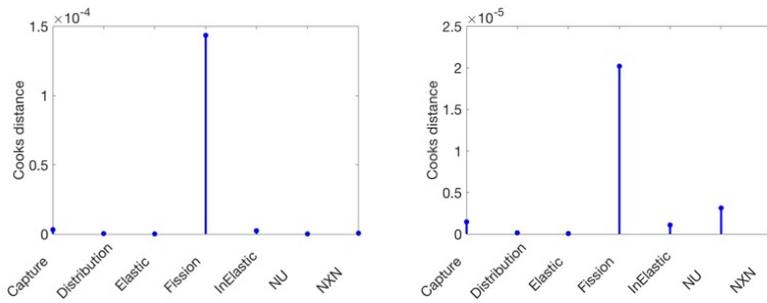


Figure 2. Cook's distance for Pu239 nuclear data (left) and Cook's distance with increased uncertainty in the prior \bar{v} (right)

It means that only the nuclear data related to Pu239 are adjusted here using the integral experiment.

The Cook's distance for this case is plotted in Figure 2 (left). The Cook's distance for the fission reaction is very large in comparison to the other nuclear reactions. It indicates that the major contribution on the adjustment is due to the k_{eff} sensitivities related to the Pu239 fission cross-section. However, from theoretical aspect and from expert opinions, it is known that for JEZEBEL experiment, both fission and \bar{v} are very influential for k_{eff} . This contradiction may suggest that there is a possibility that the prior uncertainties are not evaluated properly (too low). Thus, when the prior uncertainty on Pu239 \bar{v} data is increased by 1% and the adjustment process is repeated with JEZEBEL, the influence of Pu239 \bar{v} can now be seen on Cook's distance (see Figure 2 (right)). The Cook's distance is plotted for the case where the prior uncertainty of \bar{v} is increased by 1%. Now, it can be seen that the Cook's distance is larger for both (the fission and \bar{v}) comparing to other nuclear reactions.

Further, in the Cook's distance calculation, if Case i is the case where the i^{th} parameter is discarded in the adjustment process (see Equation 8), *i.e.* Case 1 (all data-Capture), Case 2 (all data-Distribution), Case 3 (all data-Elastic), Case 4 (all data-Fission), Case 5 (all data-Inelastic), Case 6 (all data- \bar{v}) and Case 7 (all data-nXn). In Figure 3, the trends and their uncertainties are plotted for posterior Pu239 capture cross-section for all cases. For Case 1 (where all JEZEBEL Pu239 reaction cross-sections excluding the capture are considered for the fit), no trend is visible for the adjusted values of Capture data as JEZEBEL Pu239

Capture cross-section is excluded from the adjustment. For Case 4 and Case 6, trends are very much different from the others. It indicates that the major difference is coming from fission and $\bar{\nu}$ data. Similarly, for Case 4 and Case 6, trends and uncertainties are plotted for Pu239 fission (in Figure 4) and for Pu239 $\bar{\nu}$ (in Figure 5) simultaneously. For these 2 cases also, the major contribution in the adjustment can be seen from the fission and $\bar{\nu}$. Major differences in uncertainties for some cases can be seen but it is not discussed here since the uncertainties do not make a significant contribution to the Cook's distance.

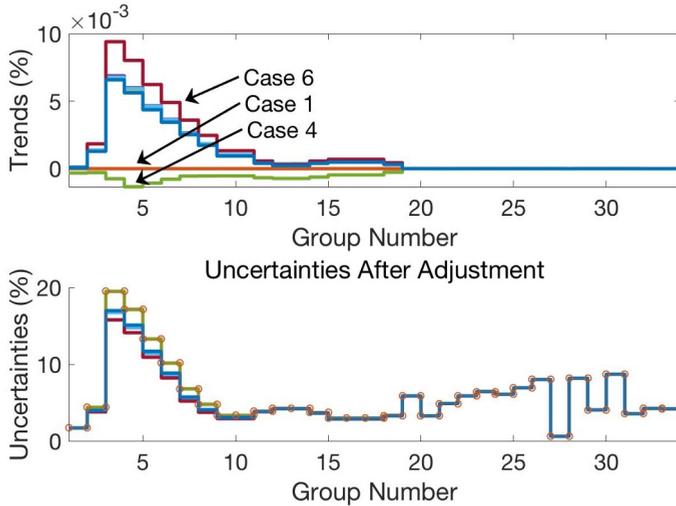


Figure 3. Adjusted Pu239 Capture data (trends and uncertainties) for all 7 sub-cases

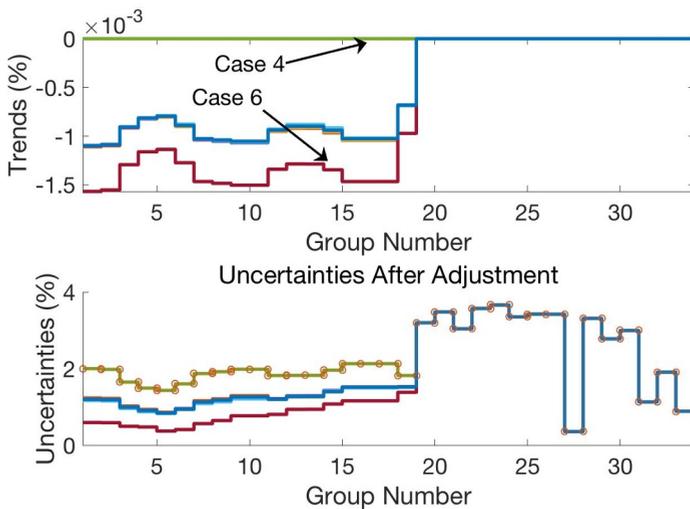


Figure 4. Adjusted Pu239 Fission data (trends and uncertainties) for all 7 sub-cases

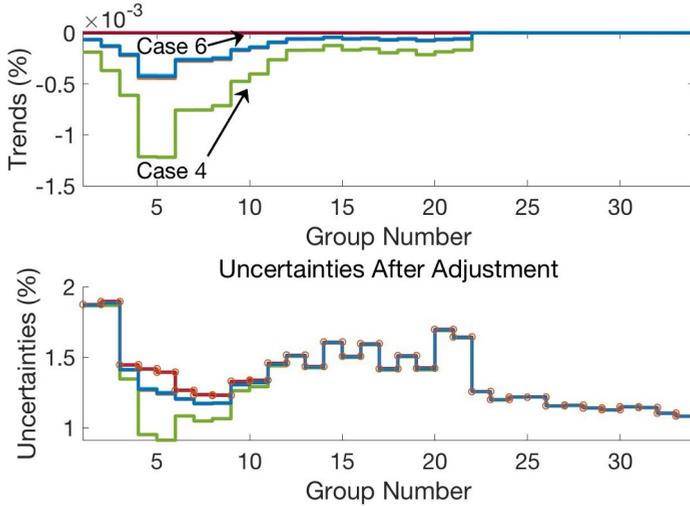


Figure 5. Adjusted Pu239 $\bar{\nu}$ data (trends and uncertainties) for all 7 sub-cases

Table 1. k_{eff} uncertainties for ASTRID

	using prior data	using prior data +1% in $\bar{\nu}$ unc.
Initial unc.	1.3227	1.3639
Final unc (fit: Pu239)	1.1365	1.1728
Individual fit effect:		
Pu239 Capture	1.32239	1.36369
Pu239 Distribution	1.31796	1.35939
Pu239 Elastic	1.32266	1.36395
Pu239 Fission	1.14785	1.19518
Pu239 InElastic	1.32245	1.36374
Pu239 Nu	1.32213	1.34499
Pu239 NXN	1.32265	1.36392

5 Transposition to ASTRID

Adjusted nuclear data using JEZEBEL integral experiment is used to compute the k_{eff} uncertainties of ASTRID GEN-IV concept reactor. A reduction in uncertainties can be seen when the posterior data is used (see Table 1). The initial uncertainty using the prior data is 1.3227% and the final uncertainty with the adjusted posterior data (when Pu239 all reaction parameters from JEZEBEL are used) is 1.1365%. It can also be seen from the Table 1 that the major reduction in the ASTRID k_{eff} uncertainty is due to the Pu239 fission reaction parameters. The final uncertainty due to Pu239 fission uncertainty is 1.14785% which is approximately equal to the final uncertainty due to Pu239 all reactions. Other reaction parameters of Pu239 isotopes are not making any significant contribution to the uncertainty reduction. However, when the prior uncertainty in $\bar{\nu}$ is increased by 1%, an additional contribution of Pu239 $\bar{\nu}$ data is also seen in the uncertainty reduction of ASTRID k_{eff} .

6 Conclusions

In this work, the influence of isotopes and nuclear reactions on adjustment is investigated using Cook's distance. For this study, the prior nuclear data uncertainties and their correlations are taken from CEA files. The JEZEBEL experiment is used as an integral experiment for nuclear data adjustment. It is concluded that the major contribution on the adjustment is due to the k_{eff} sensitivities related to the Pu239 isotope, more precisely due to Pu239 fission and Pu239 $\bar{\nu}$ reaction. The transposition of the results on ASTRID fast reactor concept is also discussed.

Acknowledgements

This study was performed in the framework of an on-going collaboration between Uppsala University and CEA. The first author gratefully acknowledges financial support from the Swedish Research Council.

References

- [1] P. Leconte et al., Feedback on Pu239 and Pu240 nuclear data and associated covariances matrix through the CERES integral experiments, *Journal of Nuclear Science and Technology*, 52:7-8, 1044-1052, 2015.
- [2] C. De Saint Jean et al., Evaluation of Neutron-induced Cross Sections and their Related Covariances with Physical Constraints, *Nuclear Data Sheets*, Volume 148, Pages 383-419, 2018.
- [3] A. Koning, S. Hilaire, S. Goriely, User Manual, Talys-1.8, A Nuclear Reaction Program, 2015.
- [4] G. Palmiotti et al., Integral Experiments Analysis for Validation and Improvement of Minor Actinide Data for Transmutation Needs, *AIP Conference Proceedings* 769, 1436, 2005.
- [5] J. Tommasi et al., Analysis of Superphenix and Phenix neutron physics experiments with the ERANOS code and data system, Operational and decommissioning experience with fast reactors. Proceedings of a technical meeting, Cadarache, France, 2002.
- [6] T. A. B. Snijders, and R. J. Bosker, *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, second edition, Sage Publishers, ISBN 9781849202008, 2012.
- [7] Methods and Issues for the Combined Use of Integral Experiments and Covariance data, A report by the working party of International Nuclear Data Evaluation Co-operation of NEA Nuclear Science Committee.
- [8] International Handbook of Evaluated Criticality Safety Benchmark Experiments, 2018.