



**HAL**  
open science

## Using the Monte-Carlo method to analyze experimental data and produce uncertainties and covariances

Greg Henning, Maëlle Kerveno, Philippe Dessagne, François Claeys, Nicolas Dari Bako, Marc Dupuis, Stephane Hilaire, Pascal Romain, Cyrille de Saint Jean, Roberto Capote, et al.

### ► To cite this version:

Greg Henning, Maëlle Kerveno, Philippe Dessagne, François Claeys, Nicolas Dari Bako, et al.. Using the Monte-Carlo method to analyze experimental data and produce uncertainties and covariances. 15th International Conference on Nuclear Data for Science and Technology (ND2022), Jul 2022, Online Conference, United States. pp.01045, 10.1051/epjconf/202328401045 . hal-04124948

**HAL Id: hal-04124948**

**<https://hal.science/hal-04124948v1>**

Submitted on 2 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Using the Monte-Carlo method to analyze experimental data and produce uncertainties and covariances.

Greg Henning<sup>1,\*</sup>, Maëlle Kerveno<sup>1</sup>, Philippe Dessagne<sup>1</sup>, François Claeys<sup>1,2</sup>, Nicolas Dari Bako<sup>1</sup>, Marc Dupuis<sup>3</sup>, Stéphane Hilaire<sup>3</sup>, Pascal Romain<sup>3</sup>, Cyrille de Saint Jean<sup>3</sup>, Roberto Capote<sup>4</sup>, Marian Boromiza<sup>5</sup>, Adina Olacel<sup>5</sup>, Alexandru Negret<sup>5</sup>, Catalin Borcea<sup>5</sup>, Arjan Plompen<sup>6</sup>, Carlos Paradela Dobarro<sup>6</sup>, Markus Nyman<sup>6,\*\*</sup>, Jean-Claude Drohé<sup>6</sup>, and Ruud Wynants<sup>6</sup>

<sup>1</sup>Université de Strasbourg, CNRS, IPHC/DRS UMR 7178, 23 Rue du Loess, F-67037 Strasbourg, France

<sup>2</sup>CEA, DES, IRESNE, DER, SPRC, LEPh, F-13108 Saint-Paul-lez-Durance, France

<sup>3</sup>CEA, DAM, DIF, F-91297 Arpajon, France

<sup>4</sup>Nuclear Data Section, International Atomic Energy Agency, Wagramer Strasse, A-1400 Vienna, Austria

<sup>5</sup>Horia Hulubei National Institute for Physics and Nuclear Engineering, 077125 Bucharest-Măgurele, Romania

<sup>6</sup>European Commission, Joint Research Centre, Retieseweg 111, B-2440 Geel, Belgium

**Abstract.** The production of useful and high-quality nuclear data requires measurements with high precision and extensive information on uncertainties and possible correlations. Analytical treatment of uncertainty propagation can become very tedious when dealing with a high number of parameters. Even worse, the production of a covariance matrix, usually needed in the evaluation process, will require lengthy and error-prone formulas. To work around these issues, we propose using random sampling techniques in the data analysis to obtain final values, uncertainties and covariances and for analyzing the sensitivity of the results to key parameters. We demonstrate this by one full analysis, one partial analysis and an analysis of the sensitivity to branching ratios in the case of  $(n,n'\gamma)$  cross section measurements.

## 1 Context

We present this paper in the context of the improvement of the evaluated nuclear data for application. This is done in order to perform better numerical simulations to optimize and predict performances, reactor control parameters and radiation safety related quantities. These databases still present large uncertainties, preventing calculations from reaching the required precision. Their improvement requires new measurements and better theoretical descriptions of involved reactions. Our primary work is focused on the measurement of  $(n, xn\gamma)$  cross-section [1–3] However, the general idea presented in this paper can be applied in many other contexts.

## 2 Motivation

When performing data analysis of experiments, many external parameters (detector efficiencies, distance of flight, ...) are involved, in order to process the raw data. Furthermore, all the steps of the data treatment (event selection, calibration, ...) may introduce uncertainties and correlations.

The usual method for combining and computing uncertainties is to use analytical developments based on the perturbation theory (e.g.  $u_{f(x)}^2 = \left(\frac{\partial f}{\partial x}\right)^2 \times u_x^2$ ). This method

works well for simple cases, but with multiple parameters and sources of uncertainty, deriving the final total combined uncertainty can be long and complex. Furthermore, it strictly applies only to small deviations from the central values, which is not always the case. Implementing the formula into the analysis code becomes a tedious process where mistakes can appear, and the final uncertainty value will be wrong.

Finally, this method makes it difficult to calculate covariances, and the inclusion of some unusual form of uncertainty (asymmetric, non-Gaussian) is not directly possible.

## 3 The Monte Carlo method

To work around all the issues presented above, we offer one possible solution: using random sampling (i.e. Monte Carlo) methods to obtain final values with their uncertainties and covariances.

The general principle is the following : Each parameter is randomly sampled according to its probability density function, all the parameters are used to compute one realization of the value of interest. This is done many times, as each computed value is stored in a *stack*. When all the iterations have been performed, the average value, standard deviation and (when applicable) covariance matrix are calculated from the *stacked* values. We note that this method allows us to turn *off* and *on* specific sources of

\*e-mail: ghenning@iphc.cnrs.fr

\*\*Currently, Univ. of Helsinki

uncertainty, by including or not the related parameters in the calculations. This makes the study of sensitivity easy, and a good way to check consistency.

The treatment of all sources of uncertainty is the same, without differentiation between the ones that come from *systematic* or *statistical* sources, as long as their probability distribution is well-defined. In fact, in this treatment, one should prefer the Type A vs. *Type B* split (rather than statistical, systematic) recommended by the *Guide to the expression of uncertainty in measurement* [4]. Particular attention should be given to choosing the probability density functions of the parameters variables. This being done, the reprocessing of data at each iteration with new parameter realizations will ensure that uncertainty from the analyzed data will be taken into account. The ability to examine the final distribution of results (and not just the central value and standard deviation) is a great benefit to check the consistency of all results given the chosen inputs.

Special attention needs to be applied to the convergence of the series. In order to produce accurate and stable final values, *enough* iterations are needed. The tricky part being that it is not evident *a priori* how many iterations is enough. One has to check the convergence of the result by inspection.

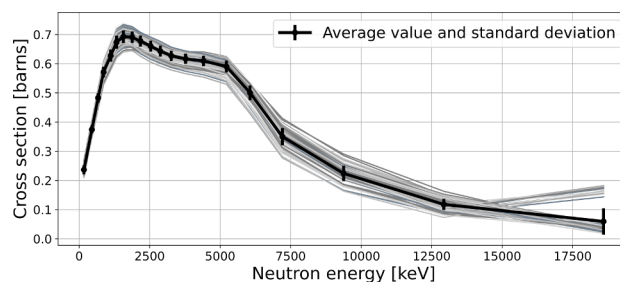
Thankfully, with modern computing infrastructures, large storage spaces and many computing units are easily available. This allows, within a reasonable time and use of computing resources, to perform a very large number of iterations.

## 4 Full Monte Carlo Analysis

The first example of applying such a technique we are presenting is the full Monte Carlo analysis. In this example, for each iteration, we will start the whole analysis pipeline (event selection, data projection, histogram fitting, ...) from scratch, with all external parameters (detector efficiency, target mass, ...) randomly sampled. The analysis is then conducted (automatically) to the end and produces one set of values (in our example,  $(n, n'\gamma)$  cross-sections for several neutron energies) which is added to the *stack*. As the spectra are re-extracted and fitted at each iteration, the procedure uncertainty, as well as uncertainties coming from the data analysis, are automatically taken into account.

When the sufficient number of iterations for convergence has been achieved (in our case, 30 is enough to reach the limit uncertainty), we use the `numpy` package [5] for Python [6] to compute, from all the values, the central values, standard deviations and covariance matrices.

Figure 1 shows an example : it is the  $^{184}\text{W}(n, n'\gamma_{111\text{ keV}})$  cross-section obtained with a full Monte Carlo analysis [7]. Each gray line in the plot represents the result of one iteration, the black points, line and error bars are the final values.



**Figure 1.**  $^{184}\text{W}(n, n'\gamma_{111\text{ keV}})$  cross-section obtained with full Monte Carlo analysis. Each gray line in the plot represents the result of one iteration, the black points, line and error bars are the final values.

## 5 Random sampling applied on intermediate results

In the cases of data that have already been analyzed (using a deterministic method), one can pick up the intermediate result files and, by applying random sampling methods, replay the last steps of the analysis many times. This is a great way to access covariance when the initial analysis did not. A similar method has been applied before in reference [8].

We applied this to  $^{238}\text{U}(n, n'\gamma)$  data [9] and the Monte-Carlo method reproduces the central value from the analytical method with a very good agreement— figure 2. The obtained uncertainties are slightly different (but still well compatible with each other). In particular, the uncertainties obtained with the Monte Carlo method present more structure than those with the analytical method. This type of structure, which is related to the input data, is not accounted for by the analytical method.

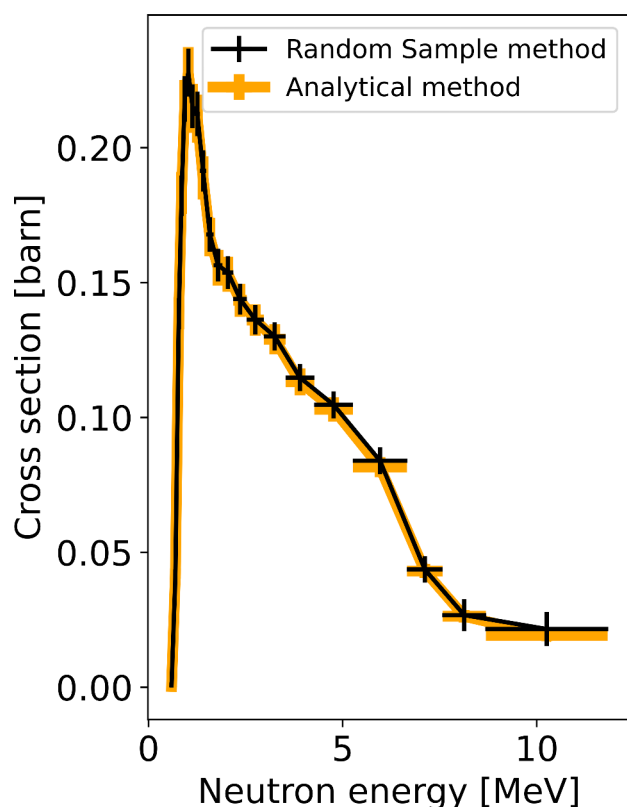
## 6 Monte Carlo method for sensitivity analysis

In addition to analyzing data using the random sampling method, the Monte Carlo method can be used to study the sensitivity to parameters. By varying a parameter through random sampling, the variation in the final results provides a sensitivity coefficient from the mean values and standard deviations of the parameter and final result.

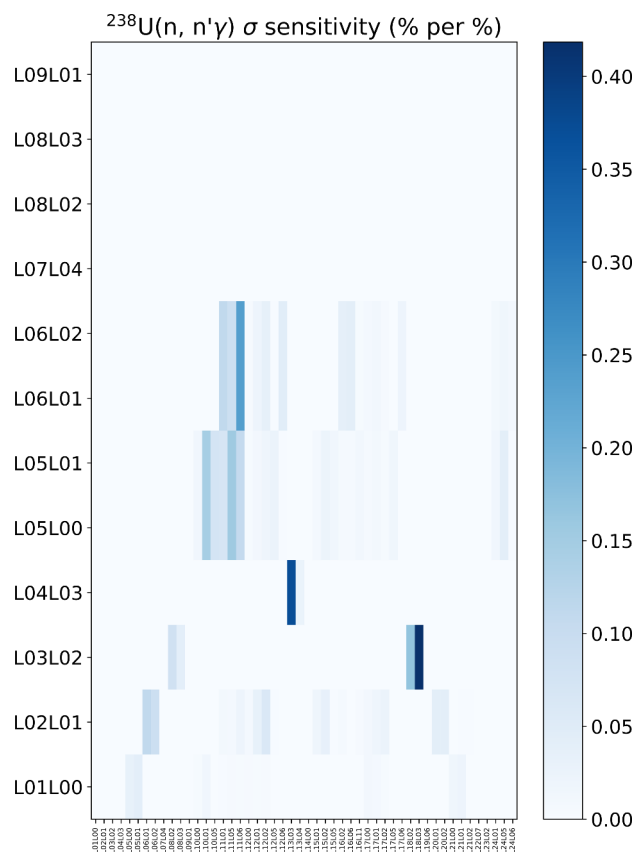
We used this method to study the sensitivity of calculated  $(n, n'\gamma)$  cross-sections to transitions branching ratios. For this purpose, we performed, for each transition in the level scheme, one hundred calculations using TALYS-1.8 [10], where the intensity of the transition is varied around its reference value [11].

After all calculations are finished, the results are loaded into a Python [6] `numpy` array [5] and the module is used to calculate the relative standard deviation expected on a calculated cross-section per relative standard deviation (i.e. uncertainty) on a specific  $\gamma$  transition branching ratio in the level scheme.

The resulting sensitivity matrix shows that some transitions have a sensitivity of as high as 40 % to other transitions' branching ratios, as seen in figure 3. See another



**Figure 2.** Cross-section for the  $^{238}\text{U}(n, n'\gamma_{653 \text{ keV}})$  reaction [9] obtained with the analytical method (yellow) and the Monte-Carlo processing of intermediate data (black).



**Figure 3.** Sensitivity matrix for transitions in the  $^{238}\text{U}(n, n'\gamma)$  reaction, in % of cross-section variation (y-axis) per % of branching ratio variation (x-axis). (Reproduced from [12].)

contribution in this conference for more details on the interpretation of the matrix [13].

## 7 Conclusion

We gave three examples of how random sampling can be used for data analysis or to study the sensitivity of calculations to parameters. It is a convenient solution to produce accurate uncertainty and covariance matrix without lengthy and error-prone uncertainty propagation. It is also a good tool to test, using calculations, the sensitivity of a value to external parameters. We believe it's a powerful technique that can be adapted in many situations and provide useful results.

## References

[1] “What can we learn from (n, xn $\gamma$ ) cross-sections about reaction mechanism and nuclear structure?”, by Kerveno, Maëlle and Dupuis, Marc and Borcea, Catalin and Boromiza, Marian and Capote, Roberto and Dessagne, Philippe and Henning, Greg and Hilaire, Stéphane and Kawano, Toshihiko and Negret, Alexandra and Nyman, Markus and Olacel, Adina and Party, Eliot and Plompen, Arjan and Romain, Pascal and Sin, Mihaela. ND 2019 : International Conference on Nuclear Data for Science and Technology (2019). 10.1051/epjconf/202023901023 <https://hal.archives-ouvertes.fr/hal-02957494>

[2] “How to produce accurate inelastic cross sections from an indirect measurement method?”, by Kerveno, Maëlle and Henning, Greg and Borcea, Catalin and Dessagne, Philippe and Dupuis, Marc and Hilaire, Stéphane and Negret, Alexandru and Nyman, Markus and Olacel, Adina and Party, Eliot and Plompen, Arjan in EPJ N - Nuclear Sciences and Technologies 4, (2018). 10.1051/epjn/2018020 <https://hal.archives-ouvertes.fr/hal-02109918>

[3] “From  $\gamma$  emissions to (n,xn) cross sections of interest : The role of GAINS and GRAPHEME in nuclear reaction modeling”, by Kerveno, M. and Bacquias, A. and Borcea, C. and Dessagne, Ph. and Henning, G. and Mihailescu, C. and Negret, A. and Nyman, M. and Olacel, A. and Plompen, M. and Rouki, C. and Rudolf, G. and Thiry, C. in European Physical Journal A 51, 12 (2015). 10.1140/epja/i2015-15167-y <https://hal.archives-ouvertes.fr/hal-02154831>

[4] Guide to the expression of uncertainty in measurement (ISO/IEC Guides 98-3) <https://www.iso.org/sites/JCGM/GUM-introduction.htm>

[5] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. Nature 585, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2. <https://doi.org/10.1038/s41586-020-2649-2>

- [6] Van Rossum, G. and Drake, F.L., 2009. Python 3 Reference Manual, Scotts Valley, CA: CreateSpace.
- [7] G. Henning, Antoine Bacquias, Catalin Borcea, Mariam Boromiza, Roberto Capote, et al.. MEASUREMENT OF 182,184,186 W (N, N'  $\gamma$ ) CROSS SECTIONS AND WHAT WE CAN LEARN FROM IT. EPJ Web of Conferences, EDP Sciences, 2021, 247, pp.09003. <10.1051/epjconf/202124709003>. <hal-0319727>
- [8] Joint Research Centre, Institute for Reference Materials and Measurements, Dessagne, P., Rudolf, G., Kerveno, M., et al., High accuracy measurements of neutron inelastic scattering cross sections : three deliverables for the ANDES project, Publications Office, 2013, <https://data.europa.eu/doi/10.2787/78760>
- [9] M. Kerveno, M. Dupuis, A. Bacquias, F. Belloni, D. Bernard, et al.. Measurement of  $^{238}\text{U}(n,n'\gamma)$  cross section data and their impact on reaction models. Physical Review C, American Physical Society, 2021, 104 (4), pp.044605. <10.1103/PhysRevC.104.044605>. <hal-03382914>
- [10] A. Koenig, S. Hilaire, M. Duijvestijn, "TALYS-1.0" International Conference on Nuclear Data for Science and Technology (2007), EDP Sciences, 2008, pp. 211-214, <http://www.talys.eu/>
- [11] Reference Input Parameter Library IAEA / Nuclear Data Service <https://www.nds.iaea.org/RIPL/3/>
- [12] Measurement of  $^{238}\text{U}(n, n' \gamma)$  cross section data and their impact on reaction models, M. Kerveno, et al., PHYSICAL REVIEW C 104, 044605 (2021)
- [13] On the need for precise nuclear structure data for high quality (n, n' $\gamma$ ) cross-section measurements. Greg Henning, M. Kerveno, Ph. Dessagne, C. Chatel, F. Claeys, N. Dari Bako, M. Dupuis, S. Hilaire, P. Romain, C. de Saint-Jean R. Capote, M. Boromiza, A. Olacel, A. Negret, C. Borcea, A.J.M. Plompen, C. Paradela, M. Nyman. Contribution to 15th International Conference on Nuclear Data for Science and Technology (ND2022).