Nuclear Archaeology Based on Measurements of Reprocessing Waste: First Experimental Results

Antonio Figueroa, Malte Göttsche
Nuclear Verification and Disarmament Group, RWTH Aachen University

Abstract

Verifying the operational history of nuclear reactors and by proxy the production of plutonium is key to nuclear disarmament efforts. For this, we develop a nuclear archaeology approach through which the history can be verified based on nuclide measurements of spent fuel and reprocessing waste. Our focus is on reconstructing burnup and cooling time values. Pursuing a probabilistic Bayesian inference framework allows us to include various sources of information beyond the measurements such as declarations and records. The goal is an ability to identify inconsistencies and to gain an insight into the actual history, including robust uncertainty estimates. This paper addresses two stages of our research: First, seeking to simultaneously take various isotopic ratios into account, we have identified those that minimize the uncertainties of our reconstruction. Since the number of theoretically possible ratios is very large, we have employed a computational technique for this purpose, Approximate Bayes Computation. Second, we have begun to test our approach initially for simple scenarios based on hypothetical simulation-based studies, but also using actual measurement data. These were obtained from the Spent Fuel Isotopic Composition 2.0 database.

Introduction

While there is extensive safeguards experience in verifying both the correctness and completeness of regularly issued nuclear material declarations issued by non-weapon states, there is a lack of methods to efficiently and directly verify nuclear material “baseline” declarations, i.e. the first verified declaration a state makes upon entering an agreement. Such methods are particularly important in the disarmament context: A solid understanding of fissile-material holdings is needed to achieve a meaningful degree of predictability and irreversibility of future arms reductions.

In addition to direct data on produced fissile materials, such records could contain information on historical operations of the nuclear facilities. For reactors, in addition to reactor and fuel designs, it could include data on reactor power, fuel burnup, and cooling time (which refers to the time passed since a specific campaign). We call these data operational parameters.

Nuclear archaeology is a toolbox that seeks to reconstruct fissile material production histories. As we have previously discussed [1], measurements of high-level reprocessing waste can be one element: This waste contains nearly all fission products and minor actinides after dissolving the spent fuel. Accordingly, it holds a rich isotopic signature of past fuel cycle activities. This waste could to some extent be used to compare results with the operational history contained in records, and as such serve to check the declaration for consistency. If a state declared that a reactor was used for civilian purposes with high burnup, this method could possibly prove that low burnup campaigns for perhaps military purposes were run. Similarly, a reactor may have run for more time than declared, which could possibly be detected by examining the cooling times. Determining the reactor power coupled with its times of operation could lead to a plutonium production estimate.

In contrast to spent fuel – for which analytical nuclear forensics methods to deduce operational reactor parameters exist, reprocessing waste contains a mixture of nuclides in different chemical phases, originated from different irradiation campaigns, produced across several years of reactor operations. Operational reactor parameters will need to be deduced numerically. While it is not clear whether the proposed approach can be applied to complex programs, it may at least be feasible for use in simpler programs of a complexity similar to the North Korean case.

We develop a method that combines measurements of selected isotope ratios, nuclear reactor models and domain information in a probabilistic framework, in order to determine plausible irradiation histories consistent with the available data. As domain information, we refer to expert knowledge and provided declarations and records (with an uncertainty attributed to it as it may be false), possibly also intelligence. The central element of our method is Bayes' theorem, which allows us to discern for instance the most likely irradiation history along with a robust uncertainty estimate.

In previous work [2], we described this method as part of a proof-of-concept study. It took, however, only five isotopes into consideration and was purely simulation-based. With this paper, we demonstrate a methodology to select a much larger set of isotopic ratios which has the potential to significantly increase our reconstruction capabilities. Furthermore, we test the approach based on actual measurement data.

Overview of Bayesian Inference

The solution to the above stated problem can be described in terms of probability distributions. Mathematically, it provides a link between the operational parameters and nuclide concentrations in terms of conditional probabilities:

$$p(x|y) \propto p(y|x) * p(x)$$

- The posterior $p(x|y)$ is the probability of the operational parameters x conditional on the measured nuclide ratios y.
- The likelihood $p(y|x)$ describes the probability of the measurement y being the result of a reactor operation described by some selection of x.
- Finally, $p(x)$ describes our prior knowledge on the past operations.

We describe the likelihood distribution with a normal distribution centered on the measured ratio and with a certain variance $\mathcal{N}(y, \sigma^2)$. In the case of several nuclide ratios being measured, and these measurements being independent of each other, the total likelihood can be written as:

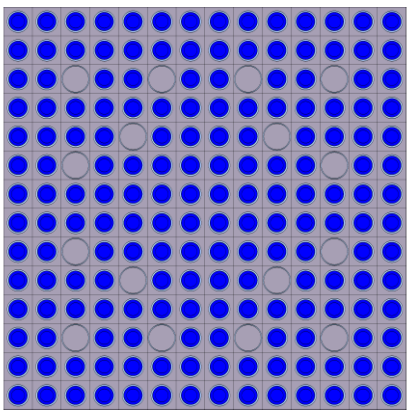
$$p(y|x) = \frac{1}{\sqrt{2\pi|\Sigma|}} e^{-\frac{1}{2}\left(Y-F(x)\right)^T * \Sigma^{-1} * \left(Y-F(x)\right)}$$

where $Y = [y_1, \ldots, y_n]$ is a vector of measurements of n nuclide ratios, $F = [f_1(x), \ldots, f_n(x)]$ is a vector of n models describing how the parameters describing the reactor operation impact the isotopic ratios of the spent fuel y_i, and $\Sigma = diag(\sigma_1^2, \ldots \sigma_n^2)$ is a n x n diagonal matrix containing the corresponding variance estimates of each measurement and model.

While elegant and compact, Bayes' theorem can be solved analytically only for simple models. In our case, a sampling method must be used in order to estimate the posterior distribution. This sampling is done via Markov Chain Monte Carlo (MCMC), a technique in which a sequence of samples is generated such that the probability of a new sample is dependent only on the probability of the previous sample [2]. We use the Python package PyMC3 [3] and its NUTS algorithm [4], which facilitates the efficient exploration of the posterior probability distribution.

Reactor Models

Bayesian inference requires a model which describes the physical system that generated the measurements, in our case, a reactor model that determines the spent fuel composition based on different input parameters. For this research, we have selected the Obrigheim reactor in Germany as abundant details on its design, as well as measurements of selected spent fuel assemblies are available in the open literature [5] and the SFCOMPO-2.0 database [6]. We have implemented a fuel assembly model using the Serpent 2 Monte Carlo software package [7]. Figure 1 shows the 2D infinite lattice model created in Serpent 2 and a table containing details of the modeled reactor.



	Obrigheim – BE170	Units
Fuel Type	Reprocessed UOX	-
Average Fuel Temperature	960.4	K
Fuel Density	10.05	g/cm3
Average Moderator Temperature	572	K
Average Moderator Density	0.7299	g/cm3

Figure 1: Left: Obrigheim PWR, 2D Fuel Assembly Model. The blue circles represent fuel rods, the grey ones control rod tubes. Right: Some details of the KKW Obrigheim reactor. Source: *[5]*

For the BE170 assembly, the SFCOMPO-2.0 database provides the irradiation history as well as measurements for a small group of nuclides using different measurement methods. To evaluate the quality of our Serpent 2 reactor model, we examined the calculated-to-experiment (C/E) ratio between the nuclide concentrations in our model and the database, following the declared irradiation history with the caveat that shutdown times were not simulated, thus creating a simplified model of the historical irradiation. Based on Figure 2 we conclude that this simplification of the irradiation history produces satisfactory results.

A typical Bayesian inference calculation requires about 10k model evaluations or simulations. To enhance the performance of the sampling algorithm, several sampling processes are run in parallel, each one generating the same number of samples. Even with powerful computers, it is neither feasible nor efficient to perform such a number of Monte Carlo reactor simulations for each reconstruction attempt. We address this through a series of surrogate models that approximate the evolution of each nuclide ratio during and after the reactor's operation, which run much faster.

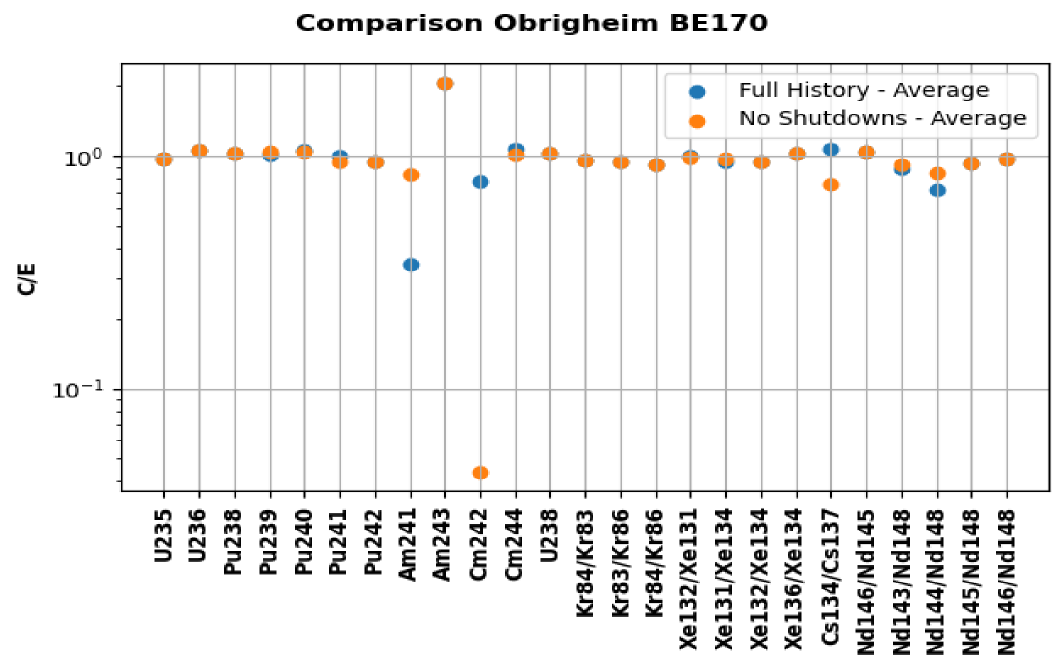


Figure 2 C/E ratio for the Obrigheim BE170 fuel assembly. We observe that the simplified model where shutdowns are omitted performs well enough for most nuclides. We have excluded from further analysis those with significant deviation (>20%) from the measured concentrations. Deviations for $^{134}Cs/^{137}Cs$ and $^{144}Nd/^{148}Nd$ are caused by the short half-lives of ^{134}Cs (2.06 years) and ^{144}Ce (284.9 days), on which ^{144}Nd depends. The major differences for ^{241}Am, ^{243}Am and ^{242}Cm are acknowledged and attributed to measurement problems by the SFCOMPO-2.0 database.

Surrogate Models

Our models are based on Gaussian Processes (GPs) [8], a technique which produces better results for this application when compared to the frequently used models based on Cubic Spline Interpolation [9]. Through a GP, a function can be interpolated without prescribing a specific function shape. Instead, a model is constructed which considers an infinite number of functions that pass between the function values and respects the described covariance matrix. As a result of this, a GP model interpolates at previously unseen input values and provides an estimate of the uncertainty of their prediction. To create each surrogate model, we have used 1000 Serpent 2 reactor simulations with burnup, power and cooling time sampled using an in-house Sobol Quasirandom Sample Generator. Such a sampling method has advantages over random and grid sampling with regard to a good coverage of the parameter space [10]. The ranges of the parameter space used for the samples is the following: power density [0.015-0.050] kW/g, burnup [0.01,60] MWd/KgHM and cooling time [1-21600] days which corresponds approximately to 60 years. The parameter ranges correspond to reasonable operational limits of the reactor. Each GP model has been made with a special Python package written in-house and uses 200 samples for training and the rest for testing the models.

Selecting optimal nuclide ratios Bayesian Inference in three steps

In our previous study we had considered only a handful of nuclides for our computer experiments. As one of the conclusions of that research, we proposed using nuclide ratios instead of individual nuclide concentrations to address the issue of complex chemical phases (supernatant, precipitates, colloids, suspensions etc.) present in reprocessing waste tanks. While the concentrations of elements might vary between phases, since isotopes of the same elements behave similarly, we would expect the isotope ratios of same elements to remain approximately constant.

We have now developed a method to determine the best-performing ratios that will enable us to reconstruct some of the operational parameters. A number of typical isotopic ratios to indicate parameters such as burnup and cooling time are well-known in the nuclear forensics field. They are usually ideal in the sense that they do not have strong dependencies on any other parameters, thus facilitating the reconstruction. Our scenarios are, however, considerably more complex, we try to deduce a larger number of parameters. Therefore, we were interested in identifying a larger set of ratios – some perhaps less obvious, as they are dependent on various parameters simultaneously. They nevertheless carry information that can be exploited by the Bayesian inference algorithm and reduce the posterior uncertainty. For this reason, we opted for a numerical selection approach.

As the measurements for the Obrigheim reactor as reported had been calculated back to the fuel discharge date, we will perform the following analysis only for burnup and reactor power. However, our approach can be repeated for other pairs of parameters, such as cooling time and power.

Our method consists of three steps. The first step is to perform a variance-based sensitivity analysis [11], which we have done using a set of 6000 reactor simulations with parameters generated from our Sobol Quasirandom Sampler. We start with 28301 possible total ratios, from which noble gases and elements with Z smaller than 31 have been excluded, as their fission yields are negligible. The sensitivity analysis allows us to filter those ratios which are neither sensitive to power nor burnup. Following this analysis, we end up with 319 ratios. From these ratios, we construct $\binom{n}{k}$= 50721 ratio pairs and generate for each one a 2D interpolator model using the SciPy library from Python [12] and 200 samples from the set of simulations.

In the second step, we compute the posterior distribution of burnup and power for all ratio pairs for different combinations of these parameters. To do this, we create two 2D grids consisting of the possible combinations of values for burnup and power. We refer to the first as the 'measurement grid', a 10x10 grid evenly spaced over the parameter ranges from which we will generate 'measurements' from the interpolator models. The second grid is 1000x1000 and we refer to it as the 'evaluation grid', as we evaluate the posterior on each point of this grid.

The evaluation of the posterior is as follows: first we choose a cell in the measurement grid and compute the 'measured value' for each ratio of the pair. We then calculate the likelihood function for each point of the evaluation grid. Figure 3 illustrates the selection method. Although straightforward, this method is, however, not practical for more than 2 or 3 parameters due to the curse of dimensionality. As the variances of the likelihood, we consider the minimum possible measurement error,[1] that is, the statistical error estimated by simple error propagation as given by [13]:

$$\frac{\sigma}{\mu}\left(\frac{N_x}{N_y}\right)[\%] = \left(\frac{1}{N_x^2}+\frac{1}{N_y^2}\right)^{-\frac{1}{2}} * 100$$

[1] As mentioned before, the error that tends to dominate in the reconstruction process is the reactor model error, however, since we don't have an error estimate for every possible ratio combination, we have decided to use the minimum statistical error instead of applying some blanket error. We have nevertheless obtained successful results with this method.

Once we calculate the posterior for each point in the evaluation grid, we can estimate the quality of the reconstruction, for this we calculate the marginal posterior distribution of each parameter and their mean and standard deviation. This is done by integrating over the grid axis opposite to the parameter of interest (e.g. to compute the marginal for power, we would have to integrate over burnup):

$$p(P|y) = \int_{P_{min}}^{P_{max}} p(B, P|y) dB$$

By repeating this process over all the points in the measurement grid, a map of the reconstruction quality for each ratio pair can be constructed over the entire range of burnup and power levels.

For the final step of our selection method, we repeat the previous process for all ratio pair combinations and rank the ratios for each cell in the measurement grid based on their maximum posterior uncertainty. From the rank we can determine a set of ratios that enable the reconstruction of burnup and power no matter their combination. While it is sufficient to take only those that optimize the posterior, it is a good practice to consider more ratios down the rank in order to make our approach more robust, thus complicating possible attempts at manipulating certain nuclide ratios.

Having applied our method, we have found a set of 54 ratios which we show partially in Table 1. While this process only guarantees our approach will work for reconstructing the history of one fuel assembly or batch of reprocessing waste, we have found that the identified optimal ratios produce very good results when reconstructing a mixture of at least two batches, as we will see in the next section.

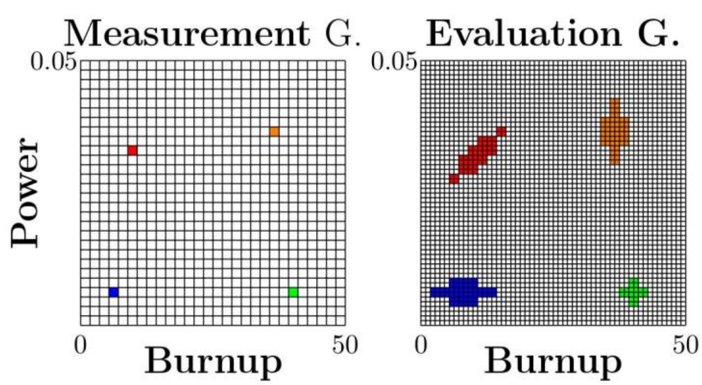


Figure 3: Ilustration of the method for selecting nuclide ratios. Since some ratios have complex dependencies on the parameters, the spread of the posterior distribution will vary for different values on the measurement grid. In the example, we observe 4 points which are mapped to the evaluation grid. The green point has the smallest posterior variance.

Table 1: Partial list of selected ratios for burnup and power

Element	Ratio	Element	Ratio	Element	Ratio
Sm	$^{148}Sm/^{152}Sm$ $^{147}Sm/^{148}Sm$	Sn	$^{116}Sn/^{122}Sn$ $^{124}Sn/^{126}Sn$	Nd	$^{145}Nd/^{146}Nd$ $^{143}Nd/^{148}Nd$
Zr	$^{90}Zr/^{92}Zr$ $^{90}Zr/^{93}Zr$	Sb	$^{121}Sb/^{125}Sb$ $^{121}Sb/^{123}Sb$	Cs	$^{133}Cs/^{134}Cs$ $^{133}Cs/^{137}Cs$
Eu	$^{151}Eu/^{154}Eu$ $^{152}Eu/^{155}Eu$	Gd	$^{154}Gd/^{155}Gd$ $^{154}Gd/^{156}Gd$	Se	$^{77}Se/^{79}Se$ $^{78}Se/^{80}Se$
Pd	$^{104}Pd/^{107}Pd$ $^{104}Pd/^{105}Pd$	Cd	$^{110}Cd/^{112}Cd$ $^{110}Cd/^{114}Cd$	Er	$^{168}Er/^{170}Er$ $^{167}Er/^{168}Er$
Ba	$^{135}Ba/^{136}Ba$ $^{136}Ba/^{137}Ba$	Dy	$^{160}Dy/^{161}Dy$ $^{160}Dy/^{162}Dy$	Sr	$^{88}Sr/^{90}Sr$

Validation and Scenario studies

Method Validation

First, we examine whether we can reconstruct information on the irradiation of the spent fuel sample as given in SFCOMPO-2.0. Therefore, we consider the measured isotopic ratios (i.e. not those just calculated) with the exception of $^{134}Cs/^{137}Cs$,$^{144}Nd/^{148}Nd$, Xe ratios, ^{241}Am, ^{243}Am and ^{242}Cm. The first two have large deviations in the C/E ratio caused by their short half-lives relative to the simplified reactor history we have considered. The Xe ratios have been omitted due to their volatile nature, as their measurement would not be possible in the case of reprocessing waste. Finally, the Am and Cm ratios have also been omitted due to their C/E ratio. Their deviations have been attributed to measurement problems by SFCOMPO-2.0 [5]. 2/3 of the nuclide ratios used for the reconstruction have C/E values between 95-105%. The rest of the ratios have values between 90-110%. The mean C/E is 98.3% with a standard deviation of 5.5%.

For the likelihood variance, we have assigned to each nuclide the C/E deviation calculated from the Serpent 2 model. Since we find that even for only 200 training samples, the GP model error is negligible and the measurement error is also very small for the nuclides we have considered, we find that the reactor model error dominates over the rest of the uncertainty sources.

Using PyMC3, we generated 40k samples from the posteriors. Unfortunately, it was not possible to reconstruct the mean power level of the reactor. This is to be expected, as based on a sensitivity analysis, we found that none of the measured ratios is strongly sensitive to power. The reconstructed burnup, however, shows a well-defined distribution whose mean matches the declared value, and whose standard deviation allows us to estimate a relative uncertainty of approximately 5%. The posterior uncertainty is smaller than the average likelihood uncertainty, which can be expected as several of the considered ratios are sensitive to burnup. This phenomenon is at the core of our Bayesian inference approach. Consequently, if we measure further informative nuclide ratios, we expect the posterior uncertainty to shrink provided new and non-redundant information is contained by the additional ratios.

Next, we attempt to reconstruct information on two different reactor operations, based on isotopic ratios of hypothetical high-level reprocessing waste, a mixture from both reprocessing campaigns.

Specifically, we have reconstructed a scenario of two batches where one corresponds to the Obrigheim spent fuel measurement, and the other is a hypothetical case. We then simulate a waste 'measurement' of the ratios present in SFCOMPO-2.0. This is done by simulating the history to be reconstructed using the GP models and obtain the isotopic ratios from them. Then, we sample each 'measurement', assuming a normal distribution with the GP results as mean value and the corresponding standard deviation estimated from the mean and the C/E value. In this case, we have simulated a low burnup campaign of 5 MWd/kg operating at the same mean power as the first case. Such a combination of batches of low and high burnup is of interest for non-proliferation, since a state could declare a civilian program with high burnup, but also run an undeclared military program at low burnup.

As can be seen in Figure 4 (left side) we can easily discern the two batches (assuming a scenario where the proportion of the mixture is 1:1), although there is considerable uncertainty on the mean values of each. It is also possible to determine the mixture proportion.

As a last scenario, we repeated this calculation but now with the nuclide ratios we identified in the previous section, obtaining both hypothetical 'measurement' values the same way as described above. This time, we assume a 5% uncertainty for each ratio. Figure 4 (right side) shows the results. The improvement is significant for both burnups and the mixture proportion. To further explore the capabilities of our method, we studied the behavior for different mixtures with different proportions. The results contained in Table 2 shows a good overall performance of the method. We observe that the absolute standard deviation of the reconstruction is approximately the same for both burnups, i.e., the relative uncertainties are larger for low values. Still, even a large standard deviation of more than 10% is well-suited to distinguish between low and high burnup campaigns. More research is needed to study posterior uncertainties for various scenarios.

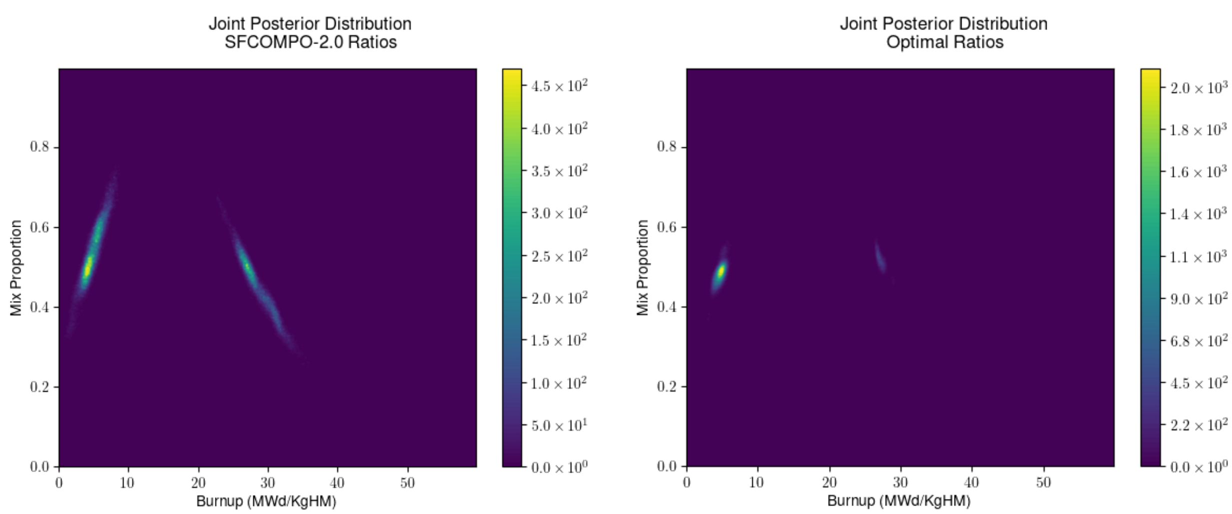


Figure 4: 2D Histogram of the Joint Posterior Distribution for the mixture of 2 batches of reprocessing waste. The left plot shows the result obtained when using the ratios from the SFCOMPO-2.0 database, the right the ones obtained using the ideal set of ratios. The brighter areas represent a larger probability. The uncertainty is much smaller when using the optimal set of ratios.

Table 2: Summary statistics for the last scenario. Reconstructed values are indicated by mean and standard deviation of the posterior distribution. In parenthesis we indicate the relative uncertainty of the reconstruction. Reconstructed burnup units are MWd/kgHM.

True Burnup 1	Reconstructed Burnup 1	True Burnup 2	Reconstructed Burnup 2	Mixture Proportion	Reconstructed Mixture Proportion
5 MWd/kgHM	5.52 +/- 0.63 (11%)	27.5 MWd/kgHM	27.83 +/- 0.73 (2.6%)	0.5	0.50 +/- 0.02 (5.3%)
	5.01 +/- 0.61 (12%)		27.45 +/- 0.61 (2.2%)	0.4	0.39 +/- 0.02 (6.7%)
	4.71 +/- 0.66 (14%)		27.20 +/- 0.33 (1.2%)	0.3	0.29 +/- 0.02 (6.7%)
	5.46 +/- 0.46 (8.5%)		28.56 +/- 1.13 (3.9%)	0.2	0.23 +/- 0.02 (11%)

Conclusion

We have demonstrated a method for the determination of an optimal set of ratios for the reconstruction of any combination of operational parameters. By design, with these ratios we achieve a very high precision and accuracy when reconstructing burnup. We have presented and successfully tested our method for the reconstructing burnup based on an actual spent fuel measurement. In addition, when considering scenarios for high-level waste consisting of two reprocessed spent fuel batches, the burnup levels of each batch and the mixing fraction can be successfully reconstructed.

While these scenarios could have possibly been solved by simpler means, taking fewer isotopes into account, our Bayesian approach has the potential to solve also more complex scenarios, such as more extensive operational histories, deducing further parameters of nuclear archaeology interest beyond burnup. Naturally, a next step will be to systematically examine the possibilities and limitations of Bayesian inference in such cases, including the use of priors.

Acknowledgements

This research has been funded by the Mathematics Division of the Center for Computational Engineering Science at RWTH Aachen University, and the Volkswagen Foundation. The funding sources had no role in the study design, collection, analysis and interpretation of data as well in the writing of the report. The calculations were run on RWTH Aachen's Compute Cluster under the rwth0572 computing project.

Bibliography

[1] M. Göttsche, A. Figueroa and M. Frank, "An Integrated Approach to Nuclear Archaeology," in *60th INMM Annual Meeting*, Palm Desert, 2019.

[2] A. Figueroa and M. Göttsche, "Nuclear archaeology: reconstructing reactor histories from reprocessing waste," *ESARDA Bulletin,* vol. 59, pp. 39-46, 2019.

[3] J. Salvatier, T. V. Wiecki and C. Fonnesbeck, "Probabilistic programming in Python using PyMC3," *PeerJ Computer Science,* 2016.

[4] A. Gelman and M. Hoffman, "The no-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo," *Journal of Machine Learning Research,* vol. 15, pp. 1593-1623, 2014.

[5] G. Radulescu, I. C. Gauld and G. Ilas, "SCALE 5.1 Predictions of Spent Nuclear Fuel Isotopic Compositions," ORNL, Oak Ridge, 2010.

[6] I. Gauld, M. Williams, F. Michel-Sendis and J. Martinez, "Integral Nuclear Data Validation Using Experimental Spent Nuclear Fuel Compositions," *Journal of Nuclear Engineering and Technology,* vol. 49.6, pp. 1226-1233, 2017.

[7] J. Leppännen, "The Serpent Monte Carlo code: Status, development and applications in 2013," *Annals of Nuclear Energy,* pp. 142-150, 2015.

[8] C. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning, Boston: MIT Press, 2006.

[9] A. Figueroa and M. Göttsche, "Gaussian Processes for Surrogate Modeling of Discharged Fuel Nuclide Compositions," *Annals of Nuclear Energy,* vol. 156, p. 108085, 2021.

[10] L. Diego, M. Pedergnana and G. Sebastian, "Smart sampling and incremental function learning for very large high dimensional data," *Neural Networks,* vol. 78, pp. 75-87, 2016.

[11] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana and S. Tarantola, Global Sensitivity Analysis: The primer, John Wiley, 2008.

[12] P. Virtanen, R. Gommers and T. Oliphant, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods,* vol. 17, pp. 261-272, 2020.

[13] A. Ulianov, O. Muntenera and U. Schaltegger, "The ICPMS signal as a Poisson process: a review of," *J. Anal. At. Spectrom,* vol. 30, pp. 1297-1391, 2015.