SCIENCE & GLOBAL SECURITY
2022, VOL. 30, NO. 1, 3–21
https://doi.org/10.1080/08929882.2022.2060599

Assessing Uncertainty in Plutonium Production Estimates Based on the Isotope Ratio Method

Benjamin Jung and Malte Göttsche

Nuclear Verification and Disarmament Group, RWTH Aachen University, Aachen, Germany

ABSTRACT
Independent estimates of lifetime plutonium production can be made using forensic measurements of characteristic indicator isotope ratios in core structural elements in shut-down nuclear reactors. Incomplete knowledge of a reactor's operational history, including fuel burnup, as well as uncertainties in nuclear cross-section data, can significantly affect such plutonium estimates, making it potentially difficult to match estimates with a state's declaration. Monte Carlo methods and sensitivity analysis techniques are used to assess the propagation of different uncertainties and their impact on plutonium estimates in infinite lattice models of a heavy-water moderated reactor (CANDU 6) and a graphite-moderated reactor (the 5 MWe reactor in North Korea), with titanium-48/titanium-49 and boron-10/boron-11 as the respective indicator isotope ratios. A tolerance interval model, with specified confidence levels, rather than one based on mean values and standard deviations, is proposed for assessing plutonium estimates based on isotope ratios measurements.

ARTICLE HISTORY
Received 28 May 2021
Accepted 23 February 2022

Introduction

A robust understanding of fissile material holdings is essential to make disarmament activities more predictable and irreversible. Speculations about unaccounted-for fissile material holdings could hinder progress because only a few kilograms of plutonium or highly-enriched uranium are necessary for building a nuclear weapon. Baseline declarations of fissile material holdings and production histories by fissile material holders are a means to increase transparency and improve trust in the disarmament process. Several precedents for such declarations exist. To be credible, the declarations need to be independently verified.[1]

One such verification tool is the Isotope Ratio Method (IRM) is, which can, in principle, be used to estimate the lifetime plutonium output of shut-down reactors. Specifically, IRM assesses the neutron fluence $\Phi = \int \phi dt$ (where ϕ is the neutron flux) in several samples taken in various

CONTACT Benjamin Jung benjamin.jung1@rwth-aachen.de Nuclear Verification and Disarmament Group, RWTH Aachen University, Schinkelstraße 2, Aachen, Germany.
© 2022 Taylor & Francis Group, LLC

locations within or very close to the core of shut-down reactors that underwent neutron activation, using isotopic measurements of trace impurities. This method was initially developed for graphite-moderated reactors—where samples from the graphite would be taken (Graphite Isotope Ratio Method, GIRM)—but can be adapted for use in other reactor designs, including heavy-water reactors. For instance, if inspectors were to return to North Korea, there would be interest in using IRM to assess samples from the graphite-moderated 5 MWe reactor in Yongbyon, which produces the North Korean plutonium.[2]

A detailed understanding of the uncertainties associated with the method is necessary to evaluate whether its results are consistent with a declaration. Although IRM can theoretically reconstruct plutonium production using only the measured sample and some basic knowledge of the reactor design, having specific information on the reactor's operating history is crucial to reducing the uncertainty of the reconstruction. Several studies have demonstrated IRM and performed some form of uncertainty assessment for specific scenarios.[3] Notably, Heasler et al.[4] found a relative uncertainty of 1.6% for a generic, graphite-moderated reactor. They identified the "reactor physics error"—uncertainties on reactor design and operational history—as the most impactful uncertainty source.

However, the abovementioned studies did not include all sources of uncertainty and they did not yield a general, versatile approach applicable to all likely scenarios.

This paper describes a Monte Carlo-based method for robust uncertainty assessment, that, in theory, can account for any uncertainty source, be it documentation-related, physics-based, or owing to other factors. The method is demonstrated in four hypothetical case studies: two different reactors (a CANDU model and the 5 MWe Yongbyon reactor), each with two different sets of uncertainties. While some of the studied uncertainty sources examined here overlap with previous studies, others have not been investigated in this context. By showing how they propagate to the plutonium estimate, this study underscores the need for a case-by-case analysis. Furthermore, this paper demonstrates a sensitivity analysis technique for identifying those uncertainty sources with a strong impact, which can then be used to systematically reduce the overall uncertainty.

The isotope ratio method (IRM)

The basics of IRM have been developed by Fetter, Gesh, and Gasner and Glaser.[5] Initially, "indicator elements" must be identified; they must be present in the permanent components of the reactor and consist of isotopes that undergo neutron activation but do not decay. Finally, they must be sensitive to the expected fluence range. A procedure to identify suitable indicator elements has been developed by de Troullioud de Lanversin and Kütt.[6]

After taking samples from the reactor and measuring the isotopic ratios, a simulation must be implemented to arrive at the plutonium estimate. This IRM simulation can be divided into three steps. The first relates an isotope ratio measurement to local fluence, the second relates local plutonium production to the local fluence. Taken together, the relationship between the isotope ratios and plutonium production is obtained. The third step extrapolates from local plutonium estimates to a global plutonium estimate for the reactor as a whole. The first two steps combine neutron transport simulations and material depletion calculations. The neutron transport simulation computes the neutronics parameters: neutron energy spectrum and one-group reaction cross-sections. These are used to solve the depletion equation, which describes the production and depletion paths of all relevant isotopes as a function of fluence—in particular those of the indicator elements and the plutonium isotopes in the fuel. The third step requires mathematical tools to infer global plutonium production from several local estimates in various locations of the reactor.

For this research, the IRM implementation is limited to the first two of the three steps: local fluence and plutonium estimates. Step 1 calculates the time evolution of the isotopic vector $\overrightarrow{N}$ of the indicator elements:

$$\overrightarrow{N}(t) = \exp(\mathbf{A}t)\overrightarrow{N}(0). \tag{1}$$

$\overrightarrow{N}$ is a vector of number densities and $\boldsymbol{A}$ is the transition matrix which depends on the neutron flux and the neutron cross-sections; decay constants are not relevant here, because the selected indicator isotopes are stable. The neutron flux is approximated with a time-averaged neutron flux $\bar{\phi}$(averaged over one irradiation cycle). This approximation is valid for low burnup and this work assumes that the reactors have been operated to produce weapons-grade plutonium, which requires low fuel burnup. It simplifies calculating the neutron fluence over a certain time period $\Phi = \int_0^t \phi(t)dt \approx \bar{\phi}t$. By dividing two components of the isotope vector and substituting $t = \Phi/\bar{\phi}$, the dimensionless isotopic ratio $R_{ij}(\Phi) = N_i(\Phi)/N_j(\Phi)$ is calculated as a function of fluence.

Step 2 computes the plutonium production Pu_0 and the fluence Φ_0 of a single irradiation cycle using a reactor simulation code (see Section Implementation) to obtain the plutonium production per unit fluence $p_0 = Pu_0/\Phi_0$.

The lifetime production can be obtained either by reconstructing the entire operating history of the reactor or by approximating the history as one or several sequences of (near) identical irradiation cycles, each sequence described by average operational parameters, e.g., average burnup of discharged fuel. For long operations with many irradiation cycles, this will be necessary to reduce complexity. In such a case, precise knowledge of each cycle's individual parameters is not necessary, as long as one can ascertain that they did not vary significantly throughout a sequence. If the true operating history deviates greatly from the simplified model, the bias needs to be accounted for.

As this work focuses on demonstrating an uncertainty assessment method to be applied in the future, rather than simulating real histories in an as-precise-as-possible manner, the simplified approach with only one sequence was used. The same method can, however, be adapted to more complex IRM models.

Everything taken together, one obtains a formula to calculate the desired isotopic ratio as a function of the reactor's lifetime plutonium production. To infer plutonium production from an isotopic ratio measurement, the above expression is inverted with numerical tools. A more thorough explanation of the implementation is provided in the supplementary material on GitHub.[7]

Sources of IRM uncertainty

Several different sources of uncertainty impact the IRM plutonium quantification. Besides the abovementioned operational parameters and potentially incomplete design information, necessary approximations in the simulation models (e.g., time discretization of depletion calculations), uncertainties of nuclear data (e.g., reaction cross-sections), and computational and measurement uncertainties contribute to the overall uncertainty. To obtain the final uncertainty interval on the plutonium estimate, these uncertainties must be quantified and propagated through the IRM simulation.[8]

When verifying a declaration, inspectors need to gather information on the operating history of the target reactor. Presumably, they would have at least some degree of access to records containing such information. In the best case, these records are detailed, complete, and fully credible. More plausible, however, are cases where some or many records were discarded or lost, or where record-keeping was insufficient—in particular in the early years. This is believed to be the case in the former Soviet Union, and the United States and the United Kingdom have also noted such challenges in their fissile material declarations.[9] Also plausible are cases where records may not be deemed trustworthy, as was the case when North Korea submitted its plutonium history to the United States.[10] Independently available information—for instance from satellite images—may help, but not necessarily fill all information gaps.

Table 1. Input uncertainty scenarios.

	YB-1	YB-2	CANDU-1	CANDU-2
Temperature (K)	–	–	343–353	333–363
Thermal power (MWt)	15–22.5	15–22.5	2000–2500	1000–2500
Burnup (MW d kg^{-1})	0.3–0.5	0.2–1.3	0.2–1.5	0.5–0.8
σ (n,γ) Ti-47[a]	–	–	±2.8%[b]	±2.8%
σ (n,γ) Ti-48	–	–	±2.8%	±2.8%
σ (n,γ) Ti-49	–	–	±2.8%	±2.8%
σ (n,α) B-10	±0.8%	±0.8%	–	–
σ (n,fis) U-238	±1.3%	±1.3%	±1.3%	±1.3%
σ (n,γ) U-238	±1.2%	±1.2%	±1.2%	±1.2%
σ (n,fis) Pu-239	±1.4%	±1.4%	±1.4%	±2.8%
σ (n,γ) Pu-239	±4.3%	±4.3%	±4.3%	±4.3%

[a] σ refers to the one-group cross-section.
[b] Interval of one relative standard deviation around the mean.

Similarly, the extent to which reactor, fuel, and target designs are known will vary. If a reactor facility is accessible for visual inspections, inspectors may be able to accurately assess relevant design parameters. However, a reactor may have used many different fuels or target designs throughout its lifetime. Inspectors would have to rely on documentation for such information. The documentation, as noted above, maybe detailed and credible, but it may also be incomplete and dubious such that design parameters may be a significant source of uncertainty.

Reactors that have not only been used for plutonium production, but also for tritium production (or other purposes) are particularly challenging to analyze—the fraction of the fluence used for each type of production must be known to obtain a reliable plutonium estimate. This challenge is not the topic of this study, but other studies have indicated that it may be possible to distinguish production modes to some extent using a variant of IRM.[11]

In Heasler et al.,[12] the largest impact on the plutonium estimate (in descending order) originated from uncertainties of the fuel pin radius, fuel temperature, graphite density, equivalent boron concentration, graphite temperature, and specific power. Compared to the reactor physics error, the impact of the other uncertainties, such as random measurement errors or contamination errors, was found to be typically around one order of magnitude lower.[13]

This paper's primary purpose is to introduce the general uncertainty assessment method, not to present a comprehensive analysis of all uncertainty sources. Therefore, the following analysis focuses on average burnup, specific power and moderator temperature (operational parameters), and neutron cross-sections of boron and titanium isotopes, as well as uranium-238 and plutonium-239 (see Table 1). Other uncertainty sources are ignored, not because they are not influential—Heasler et al. have shown that they are—but to reduce the complexity and computational cost of the analysis.

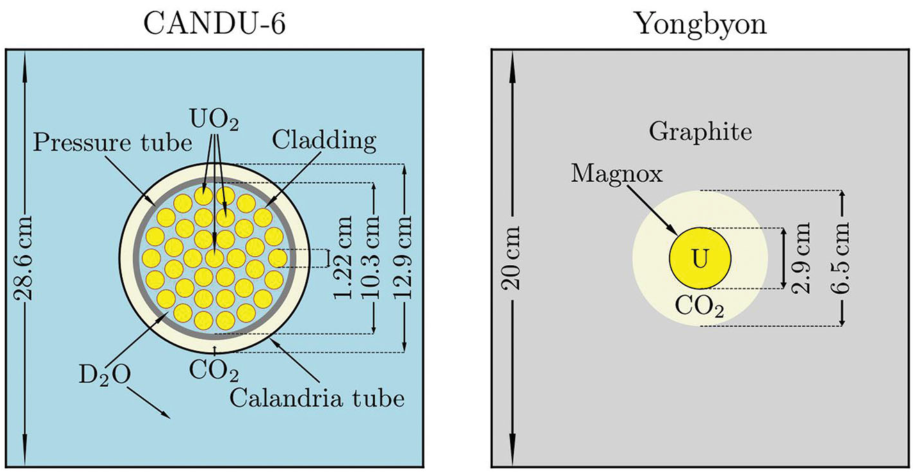

Figure 1. Unit cells of the CANDU reactor on the left and the Yongbyon reactor on the right, as implemented in Serpent 2.

Implementation

Reactor simulations

The neutron transport and the fuel depletion are simulated with Serpent 2,[14] a validated continuous-energy Monte-Carlo reactor physics code. Infinite lattice models of two reactors were implemented, which emulate a fuel channel in the center of the reactor cores. This analysis uses two different reactor models: a heavy-water moderated CANDU 6 model and the graphite-moderated 5 MWe "Yongbyon" reactor. The former is relevant because its design is similar to other heavy-water reactors used for military plutonium production (e.g., the Indian CIRUS reactor). The latter is relevant because it is the primary plutonium source in North Korea.

Figure 1 shows the fuel cell designs.[15] The implemented models were validated by comparing the spent fuel composition and the estimated plutonium production rate, respectively, for CANDU and Yongbyon with existing data and literature (see Appendix). Furthermore, the neutron spectra in the first and the last step of a simulation were compared (see Figures 7 and 8 in the Appendix). The variation between the spectra is low enough to justify approximating an average spectrum for each simulation.

This study uses boron-10/boron-11 and titanium-48/titanium-49 as indicators, respectively, for the Yongbyon and CANDU reactors. Both ratios have been identified as suitable indicators in earlier stages of IRM development. Boron is suitable in a low fluence range (due to the high neutron capture cross-section of boron-10) and is present as impurities in reactor graphite. Therefore, it is a suitable indicator for the Yongbyon reactor. In a CANDU reactor, much higher fluence ranges are to be expected. Titanium is a suitable indicator at high fluences and is present as impurities in the calandria tubes of CANDU reactors, which are made of a zirconium alloy. It is also present in aluminum alloys, which are typical materials for pressure tubes in non-power reactors.[16]

Parametrizing the uncertainties

The first step of the proposed uncertainty assessment framework is to quantify the uncertainty sources. Different sources can follow different probability distributions, which should be accounted for. The individual source uncertainties of the four case studies are shown in Table 1.

Some information on the operational history of the North Korean 5 MWe reactor is known from independent analyses. The range of the reactor's thermal power (scenarios YB-1 and YB-2) is based on Albright and O'Neill.[17] Due to the low power density, its effect on the graphite moderator temperature is neglected. It is also roughly known when the reactor was operational and for how long; de Troullioud de Lanversin and Kütt[18] provide burnup estimates for each irradiation cycle. The burnup range in scenario YB-1 is loosely derived from their work. To demonstrate the impact of the magnitude of source uncertainties, the burnup uncertainty of scenario YB-2 is not based on actual knowledge of individual cycles. Instead, it represents a burnup range extending beyond what has been observed in North Korea but could nevertheless be achieved in the future. At the upper limit of 1.3 megawatt-days per kilogram (MW d kg^{-1}), the plutonium consists of roughly 90% plutonium-239, which may still be useful for constructing nuclear weapons.

The operational parameter uncertainties of the CANDU-1 scenario are loosely derived from reference values of the CANDU 6 model.[19] Since it is assumed that the reactor has been used to produce plutonium for a military program, the upper limit on burnup is chosen to be lower than the typical burnup of CANDU fuel. Similar to the two Yongbyon scenarios, the two CANDU scenarios differ in terms of knowledge of the individual parameters.

The parameters of operational history are described by uniform probability distributions—assuming inspectors do not know one most likely (mean) value, but only a plausible range of values, all of which are equally possible. How to determine such a range in practice depends on the available information. Inspectors would need to analyze operating records as well as consult other sources of information. The ranges in Table 1 are not based on an in-depth analysis. Instead, they are loosely based on publicly available information and constructed in such a way as to showcase several different possibilities.

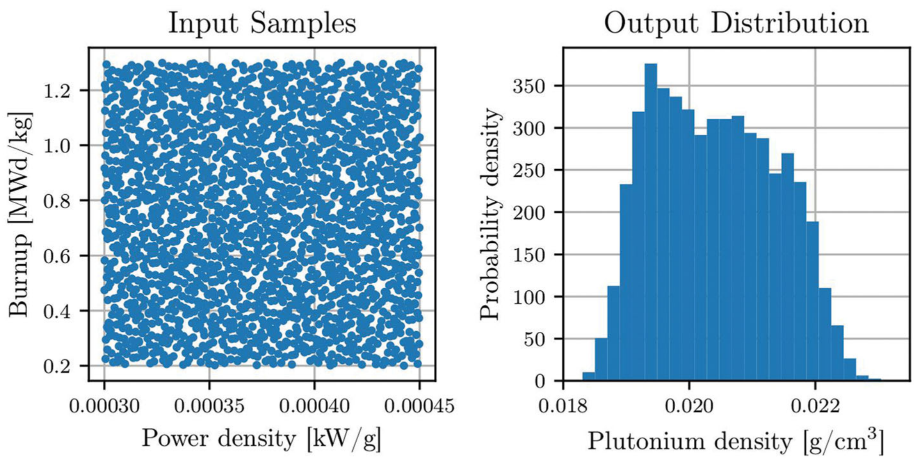

Figure 2. An example of input and output in the YB-2 scenario. The left graph shows a 2D projection of the input parameter samples (note that all parameters from Table 1 were varied). The graph on the right shows the output distribution for the plutonium estimate. In this example, the ratio boron-10/boron-11 is 0.04 and the fluence Φ is $\sim 1.46 \times 10^{21}$ cm^{-2}.

The cross-sections, however, are based on physical measurements, and their uncertainties are therefore best described by a normal distribution. Energy-dependent cross-section uncertainties from ENDF/B-VIII are used, together with time-averaged neutron spectra of typical reactor operations, to estimate a relative uncertainty for each one-group cross-section.[20]

Monte Carlo-based uncertainty propagation

The second step of the uncertainty assessment is the propagation of the uncertainty sources to the plutonium estimate. For this purpose, this paper proposes to use a (quasi-) Monte Carlo approach. Such a numerical method is adequate, as an analytical error propagation will in many cases not be feasible due to the complex impact of some of the source uncertainties. Samples are drawn from the probability distributions of the input parameters.[21] For each sample, the output (the reconstructed plutonium production) is computed. The result of the Monte Carlo approach is a probability distribution for the output. Each of the columns of Table 1 serves as a 7- or 10-dimensional input parameter space. Figure 2 illustrates the Monte Carlo approach for the YB-2 scenario. Given a hypothetical measurement boron-10/boron-11 = 0.04 and a set of parameters, the IRM model predicts a plutonium density. The parameters are varied by sampling values from the associated probability distribution (Figure 2, left) and the IRM prediction is repeated with each different set of parameter values. Accordingly, the predicted plutonium densities also vary and, when turned into a histogram, produce the graph on the right of Figure 2.

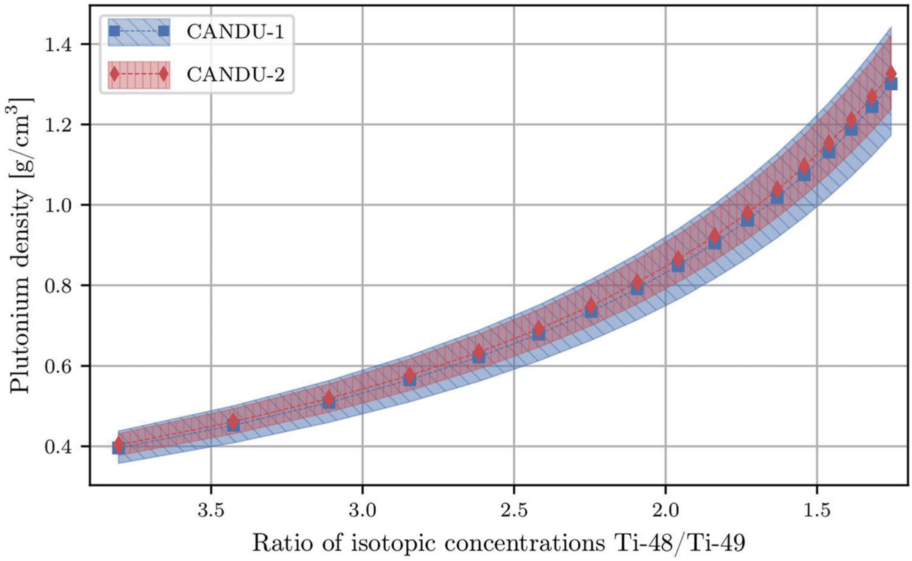

Figure 3. Reconstructed plutonium density with the CANDU reactor model. The shaded area indicates the 95% tolerance interval. The relative uncertainty of CANDU-1 is 10% and of CANDU-2 is 7%.

A meaningful description of IRM uncertainties

The third and final step is a description of the uncertainty of the final plutonium estimate. The GIRM implementation by Heasler et al.[22] already includes an uncertainty treatment; they derive a best-guess total plutonium estimate with a standard error. In some cases, however, such an estimate resulting from best-guess values of uncertain parameters may be inaccurate and misleading.

Instead, this paper proposes to take uncertainties into account by reconstructing past plutonium production in the form of tolerance intervals, which indicate the range of plutonium estimates consistent with the IRM measurement and analysis. A declaration can be accepted or rejected by assessing whether the declared plutonium inventory lies within this interval.

Furthermore, tolerance intervals are also useful in cases where the underlying probability distribution for plutonium production is non-normal. As Figure 2 shows, this may well be the case, especially when some of the uncertainty sources (particularly those related to operational history) are best described by uniform distributions within a range. In such a case, a declaration cannot be robustly tested based on a mean value and standard deviation.

Mathematically speaking, a tolerance interval contains a certain proportion p of the population with a confidence level γ. It can be computed for

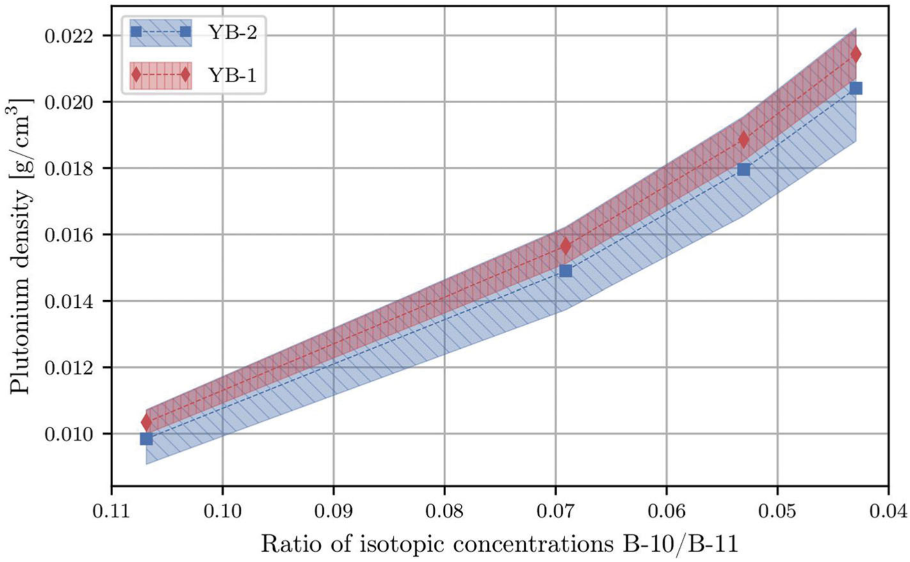

Figure 4. Reconstructed plutonium density estimates with the Yongbyon reactor model. The shaded area indicates a 95% tolerance interval. The relative uncertainty of YB-1 is 3.4% and of YB-2 is 8.4%.

non-parametric distributions using order statistics, which is the approach used in this study.[23]

This work presents p=95% tolerance intervals with a γ=95% confidence level. In the context of reconstructing plutonium production, this means that 95% of the plutonium values are consistent with the IRM analysis and the measured isotopic ratios lie within the bounds of the interval. The confidence level states that in 95 out of 100 cases, repeating the analysis will yield the same results.

Case study results

Computed tolerance intervals

The results of the uncertainty analysis of the four case studies are shown in Figures 3 and 4, respectively, for CANDU and Yongbyon. Each graph shows statistical tolerance intervals of the plutonium estimate for a range of isotopic ratio values. These isotopic ratios represent hypothetical measurements, and the ranges of displayed values correspond to the expected fluency in both reactors.[24] The plutonium production estimates are local estimates and as such quantified in units of grams per cubic centimeter (g cm^{-3}).

Several features stand out in both graphs. First, the difference between the widths of the tolerance intervals is significant. Since the cross-section uncertainties are the same for all scenarios, the difference originates

Table 2. Plutonium is estimated in tolerance intervals (p=95%,γ=95%) for a hypothetical measurement.

Scenario	Fluence (cm^{-2})	Indicator	Ratio	Pu (kg)
YB-1	1.46×10^{21}	B-10/B-11	0.04	57–62
YB-2				51–62
CANDU-1	1.14×10^{23}	Ti-48/Ti-49	1.7	8415–10334
CANDU-2				8879–10184

Plutonium densities are scaled with the reactor fuel volume.

from the operational parameters. Later, the sensitivity analysis will explain which parameters, in particular, are responsible for the difference.

Second, the tolerance intervals of the scenarios YB-1 and CANDU-2 lie entirely within those of YB-2 and CANDU-1, although their respective mean values are different. This observation is an example of why tolerance intervals are more suitable for verifying plutonium declarations than mean values with standard deviations. The former accepts all declared values within the reconstructed interval, whereas the latter prefers values closer to the mean. The tolerance interval also informs about the amount of material potentially not accounted for in the declaration.

Third, the relative uncertainty—the width of the tolerance interval relative to the mean of the interval—is constant over the range is isotopic ratios shown in each graph, meaning that the absolute uncertainty increases with the amount of plutonium produced. In a disarmament verification context, the reconstruction uncertainty on large stockpiles will be larger than on small stockpiles. While this result is not unsurprising, it means that more weapon-equivalents worth of plutonium is potentially unaccounted for.

Table 2 shows one tolerance interval of each scenario, scaled with the volume of the fuel in the reactor. Such simple global estimates do not take the spatial inhomogeneity of the neutron flux into account and, therefore, include additional bias. However, they do serve to illustrate the possible extent of the total uncertainty. Even when only considering a small set of uncertainty parameters, the relative uncertainties range from 3.4 to 10%. In the Yongbyon scenarios, these scaled tolerance intervals correspond to roughly one and three weapon-equivalents of plutonium (assuming one weapon-equivalent is 4 kg). In the CANDU scenarios, which have a higher total fluence, the scaled tolerance intervals correspond to several tons of plutonium.

Sensitivity analysis

Variance-based sensitivity analysis decomposes the variance of the model output into fractions, called global sensitivity indices, that are attributed to

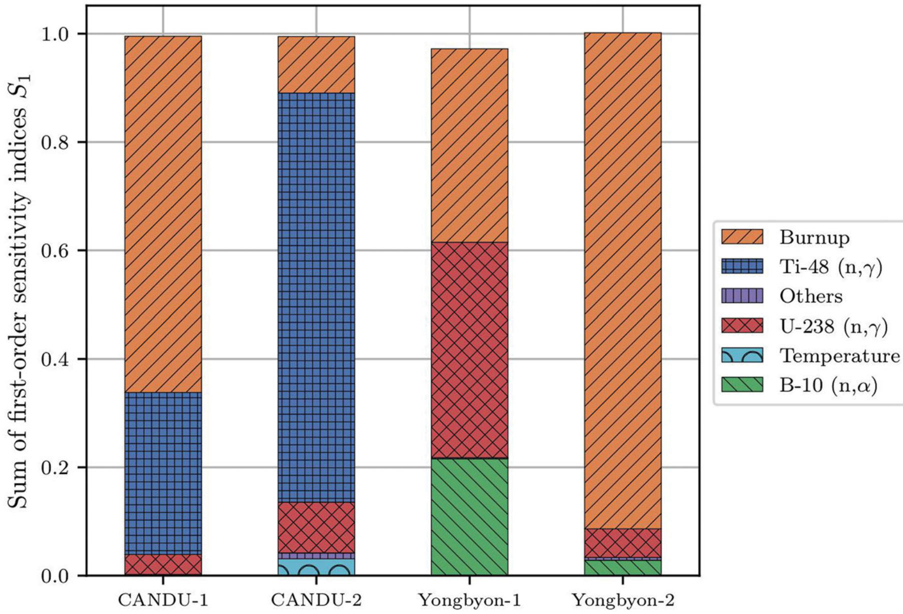

Figure 5. First-order sensitivity indices of the four scenarios. The bars are shaded according to the contribution of individual parameters. Parameters with an index close to zero are grouped into the category "others" for display purposes.

individual input parameters (first-order) or combinations of input parameters (higher-order). First-order indices describe the direct effect of a parameter on the output variance. It indicates by how much the variance would, on average, be reduced, if the parameter were to be fixed. Higher-order indices describe the effect of a combination of parameters that cannot be explained by the sum of their individual direct effects. The sum of all indices of one parameter is called the total index. If it is zero, the associated parameter can be fixed to any value within its range, without significantly altering the output variance. The sum of first- and higher-order indices of all parameters equal one by definition. This property is useful for reducing the computational cost of sensitivity analysis. If the sum of first-order indices is close to one, calculating second- or higher-order indices can be omitted, because their effect will be small.[25]

Since the values of sensitivity indices depend on the variation of the model input and output, each scenario is analyzed separately. Figure 5 displays the first-order sensitivity indices of the uncertainty estimate in Table 2. Burnup is the most influential parameter in two of the four scenarios, followed by cross-section uncertainties of the (n,γ) reaction in titanium-48, uranium-238 as well as the (n,α) reaction in boron-10. In the CANDU-2 scenario, the moderator temperature also plays a non-negligible role. The indices corresponding to the other points in Figures 3 and 4 do not differ significantly. In three of the four scenarios, the first-order indices sum to approximately one, indicating a negligible impact of higher-order effects. Only in the YB-1 scenario, which has the overall lowest uncertainty, do higher-order effects begin to become relevant.

The sensitivity analysis helps interpret the results of the uncertainty analysis. It explains why the output uncertainty of CANDU-1 is higher than that of CANDU-2. The burnup uncertainty in CANDU-1 is much higher than in CANDU-2. Conversely, the power uncertainty is much higher in CANDU-2 than in CANDU-1. The first-order index of power is approximately zero in both scenarios (see "others" in Figure 5), indicating that power has next to no influence on the output variance. The first-order index of burnup is 0.66 ± 0.03 in CANDU-1. The reduced burnup uncertainty in CANDU-2 leads to reduced output uncertainty compared to CANDU-1.

Similarly, the sensitivity analysis of the scenarios YB-1 and YB-2 explains the results of the uncertainty analysis. Reducing the uncertainty of the burnup parameter reduces the uncertainty of the plutonium estimate significantly. Furthermore, the Yongbyon scenarios exemplify how sensitivity analysis can be useful for IRM applications. In this example, YB-2 represents the inspectors' initial analysis. The sensitivity analysis indicates burnup is responsible for most of the uncertainty of the plutonium estimate. The inspectors then pursue a targeted investigation to gain more precise information on this parameter. The new information leads to a new input uncertainty assessment, namely YB-1, which leads to a plutonium estimate with less uncertainty.

In general, sensitivity analysis is model-specific and one must be cautious before generalizing the results. Nevertheless, burnup has a significant effect on the uncertainty of the plutonium estimate in all four scenarios of this study, which is remarkable because it has not been considered an uncertainty parameter in previous studies. The sensitivity analysis also shows that uncertainties in nuclear cross-section data are not negligible and can have as much impact as some operational parameters. In particular, the (n,γ) cross-sections of uranium-238, titanium-48, and boron-10 have an impact on the plutonium estimate.

Conclusion

IRM uses forensic measurements of isotopic ratios and simulation-based analysis to estimate the amount of plutonium produced in a nuclear reactor. Such an analysis needs to consider uncertainties of the parameters of the simulation. The present study demonstrated a Monte Carlo-based approach to uncertainty propagation using fictitious case studies. This approach uses prior information to define a probability distribution for each parameter and draws quasi-random samples from these distributions. The plutonium estimation procedure is applied to each sample and the resulting probability distribution is characterized to obtain a tolerance interval. Such a tolerance interval of the total plutonium production is suitable for assessing a declaration.

Despite only considering a subset of all uncertain simulation parameters, this study has shown that tolerance intervals can be large (≈10%). The sensitivity analysis has indicated that average burnup and cross-section data of uranium-238, titanium-48, and boron-10 account for most of the uncertainty in the case studies. These parameters have not been identified as significant sources of uncertainty in previous studies of (G)IRM. Sensitivity analysis adds value to IRM as a tool for verifying fissile material declarations. Inspectors could pursue a targeted attempt to obtain further information on a parameter that has been identified as highly impactful by using additional technical verification methods or by demanding such information from the inspected state.

This study has been limited to a simplified IRM implementation and hypothetical cases. Future research should examine uncertainties of global estimates, making use of full-core instead of infinite lattice reactor simulations, and include a more detailed IRM implementation that can take into account changes in operational parameters over the lifetime of a reactor. Lastly, the combined impact of all plausible uncertainty sources beyond burnup and cross-section data should be assessed.

Acknowledgments

The authors thank Jakob Brochhaus for his support with the sensitivity analysis.

Author contributions

Benjamin Jung: methodology, software, investigation, writing–original draft, writing–review, and editing. Malte Göttsche: conceptualization, methodology, writing–original draft, writing–review, and editing.

Disclosure statement

The authors have no conflict of interest to declare.

Notes and References

1. Alexander Glaser and Malte Göttsche, "Fissile Material Stockpile Declarations and Cooperative Nuclear Archaeology," in *Verifiable Declarations of Fissile Material Stocks: Challenges and Solutions* (United Nations Institute for Disarmament Research, 2017), 25–38.

2. Steve Fetter, "Nuclear Archaeology: Verifying Declarations of Fissile-Material Production," *Science & Global Security*, 3 (1993): 237–259; Christopher J. Gesh, *A Graphite Isotope Ratio Method Primer: A Method for Estimating Plutonium Production in Graphite Moderated Reactors*, PNNL-14568 (Richland, WA: Pacific Northwest National Laboratory, 2004), https://doi.org/10.2172/15010619; Alex Gasner and Alexander Glaser, "Nuclear Archaeology for Heavy-Water-Moderated Plutonium Production Reactors," *Science & Global Security*, 19 (2011): 223–233; International Panel on Fissile Materials, *Global Fissile Material Report 2009* (2009); Jungmin Kang, "Using the Graphite Isotope Ratio Method to Verify the DPRK's Plutonium-Production Declaration," *Science & Global Security*, 19 (2011): 121–129, https://doi.org/10.1080/08929882.2011.586309.
3. Gasner and Glaser, Nuclear Archaeology, op. cit.; Kang, Using the Graphite Isotope, op. cit.; Julien de Troullioud de Lanversin and Moritz Kütt, "Verifying North Korea's Plutonium Production with Nuclear Archaeology," *Science & Global Security*, 29 (2021): 145–166, https://doi.org/10.1080/08929882.2021.1999054.
4. Patrick G. Heasler et al., "Estimation Procedures and Error Analysis for Inferring the Total Plutonium (Pu) Produced by a Graphite-Moderated Reactor," *Reliability Engineering & System Safety*, 91 (2006): 1406–1413.
5. Fetter, Nuclear Archaeology, op. cit.; Gesh, A Graphite Isotope, op. cit.; Gasner and Glaser, Nuclear Archaeology, op. cit.
6. de Troullioud de Lanversin and Kütt, Verifying North Korea's Plutonium, op. cit.
7. Benjamin Jung and Malte Göttsche, "Supplementary Material," https://github.com/jung-benjamin/irm-paper-sgs.
8. Moritz Kütt, Julien de Troullioud de Lanversin, and Malte Göttsche, "Understanding Uncertainties in Nuclear Archaeology," in *59th Annual INMM meeting, Baltimore, MD* (2018), 22–26.
9. U.S. Department of Energy, *Plutonium: The First 50 Years*, DOE/DP-0137 (1996), https://doi.org/10.2172/219368; U.K. Ministry of Defense, *Plutonium & Aldermaston: An Historical Account* (2000).
10. Helene Cooper, "In Disclosure, North Korea Contradicts U.S. Intelligence on Its Plutonium Program," (2008), https://www.nytimes.com/2008/05/31/world/asia/31korea.html.
11. Julien de Troullioud de Lanversin, Malte Göttsche, and Alexander Glaser, "Nuclear Archaeology to Distinguish Plutonium and Tritium Production Modes in Heavy-Water Reactors," *Science & Global Security*, 26 (2018): 70–90.
12. Heasler et al., Estimation Procedures, op. cit.
13. It should be noted that the impact of measurement errors is small only if the indicator elements are chosen such that they are well-suited for the expected fluence range (Gasner and Glaser, Nuclear Archaeology, op. cit.; de Troullioud de Lanversin and Kütt, Verifying North Korea's Plutonium, op. cit.).
14. Jaakko Leppänen et al., "The Serpent Monte Carlo Code: Status, Development and Applications in 2013," *Annals of Nuclear Energy*, 82 (2015): 142–150.
15. The CANDU design is based on Robin Chaplin, "Genealogy of CANDU Reactors," *in The Essential CANDU, A textbook on the CANDU Nuclear Power Plant Technology*, ed. W. J. Garland, https://www.unene.ca/education/candu-textbook. The Yongbyon implementation is based on David Albright and Kevin O'Neill, *Solving the North Korean Nuclear Puzzle* (Institute for Science and International Security, 2000); Kang, Using the Graphite Isotope, op. cit.
16. Gesh, A Graphite Isotope, op. cit.; Gasner and Glaser, Nuclear Archaeology, op. cit.

17. Albright and O'Neill, Solving the North Korean, op. cit.
18. de Troullioud de Lanversin and Kütt, Verifying North Korea's Plutonium, op. cit.
19. Chaplin, Genealogy of CANDU Reactors, op. cit.
20. One-group cross-sections are obtained by weighting the energy dependent cross-sections with the neutron ux spectrum. JANIS 4.0 provides numerous nuclear data libraries in a processed format (N. Soppera, M. Bossant, and E. Dupont, "JANIS 4: An Improved Version of the NEA Java-based Nuclear Data Information System," *Nuclear Data Sheets*, 120 (2014): 294–296, https://doi.org/10.1016/j.nds.2014.07.071). Among others, it provides ENDF/B-VIII cross-section data on a specific energy grid with a covariance matrix, which was used to approximate the one-group cross-section uncertainties (D. A. Brown et al., "ENDF/B-VIII.0: The 8th Major Release of the Nuclear Reaction Data Library with CIELO-project Cross Sections, New Standards and Thermal Scattering Data," *Nuclear Data Sheets*, 148 (2018): 1–142, https://doi.org/10.1016/j.nds.2018.02.001).
21. Input samples are drawn with the Sobol sequence. Being a low-discrepancy sequence, it offers the advantage that it covers the multidimensional parameter space more evenly and with a lower number of samples compared to pseudo-random sampling methods. The Sobol sequence generation algorithm is based on the implementation described in Paul Bratley and Bennett L. Fox, "Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator," *ACM Transactions on Mathematical Software (TOMS)*, 14 (1988): 88–100. It requires initialization numbers as starting points for the sequence. The initialization numbers were identified in Stephen Joe and Frances Y. Kuo, "Constructing Sobol Sequences with Better Two-Dimensional Projections," *SIAM Journal on Scientific Computing*, 30 (2008): 2635–2654 to confer good low-discrepancy properties in high dimensions.
22. Heasler et al., Estimation Procedures, op. cit.
23. K. Krishnamoorthy and T. Mathew, *Statistical Tolerance Regions: Theory, Applications, and Computation* (John Wiley & Sons, 2008).
24. The upper limit of the fluence range for the CANDU reactor is based on an approximation of the total fluence in a CANDU reactor with a lifetime power production of the Bruce 1 reactor in Canada. The limit for the Yongbyon reactor is based on a re-simulation of the entire reactor operation following the analysis of Hecker, Braun, and Lawrence, *North Korea's Stockpiles of Fissile Material.*
25. A mathematical derivation of global sensitivity indices can be found in Ilya M. Sobol, "Global Sensitivity Indices for Nonlinear Mathematical Models and their Monte Carlo Estimates," *Mathematics and Computers in Simulation*, 55 (2001): 271–280. The first-order and total indices were computed with the Python package described in Jon Herman and Will Usher, "SALib: An Open-Source Python Library for Sensitivity Analysis," *The Journal of Open Source Software* (2017), https://doi.org/10.21105/joss.00097. The algorithm implemented in the package is based on Andrea Saltelli et al., "Variance based Sensitivity Analysis of Model Output. Design and Estimator for the Total Sensitivity Index," *Computer Physics Communications*, 181 (2010): 2, 259–270, which studies different methods for computing sensitivity indices. To estimate the uncertainty of the sensitivity indices empirical bootstrapping methods (Andrea Saltelli et al., *Global Sensitivity Analysis: The Primer* (John Wiley & Sons, 2008)) were used and the convergence criteria presented in Fanny Sarrazin, Francesca Pianosi, and Thorsten Wagener, "Global Sensitivity Analysis of Environmental Models: Convergence and Validation," *Environmental Modeling & Software*, 79 (2016):

135–152 were applied to ensure the ranking and the screening process is statistically robust.

26. Franco Michel-Sendis et al., "SFCOMPO-2.0: An OECD NEA Database of Spent Nuclear Fuel Isotopic Assays, Reactor Design Specifications, and Operating Data," *Annals of Nuclear Energy*, 110 (2017): 779–788.
27. Albright and O'Neill, Solving the North Korean, op. cit.; Kang, Using the Graphite Isotope, op. cit.; Hecker, Braun, and Lawrence, North Korea's Stockpiles, op. cit.

Funding

This research has been funded by a Freigeist Fellowship grant of the VolkswagenStiftung. The simulations were performed with computing resources granted by RWTH Aachen University under projects rwth0504, rwth0572, and thes0669.

Appendix

To gain greater confidence in the implementation of the reactor simulations, the results were compared to available data and literature. For the CANDU reactor, the burnup history of a fuel assembly was resimulated, for which measurement values of uranium and plutonium isotope concentrations are available in the SFCOMPO-2.0 database.[26] Except for the concentration of plutonium-241, the results are in good agreement (see Table 3). The most likely reason for the discrepancy in plutonium-241 is its low half-life (≈14 years). The database does not list the time between the last irradiation period of the fuel and when the isotopic composition was analyzed. Therefore, the decay of plutonium-241 is not accurately accounted for in the simulation.

For the Yongbyon reactor, plutonium production was compared with the results of several studies.[27] Figure 6 shows that the results are roughly in agreement. The reason for the discrepancies cannot be determined without more detailed knowledge about model assumptions in the other publications. For the purpose of this study, it is sufficient to see that the trend agrees.

Table 3. Isotopic concentrations of uranium and plutonium in the spent fuel of the Bruce A reactor.

	Simulation		Measurement		
Isotope	Average	Std error	Average	Std error	Deviation in #std
U235/U	0.245	0.058	0.241	0.004	0.076
U236/U	0.072	0.008	0.081	0.004	1.101
U238/U	99.683	0.050	99.671	0.010	0.230
Pu239/Pu	69.126	3.418	68.788	0.159	0.099
Pu240/Pu	26.757	2.291	25.147	0.106	0.702
Pu241/Pu	2.563	0.480	4.456	0.089	3.879
Pu242/Pu	1.451	0.648	1.249	0.057	0.309

Measurement values are taken from SFCOMPO-2.0.

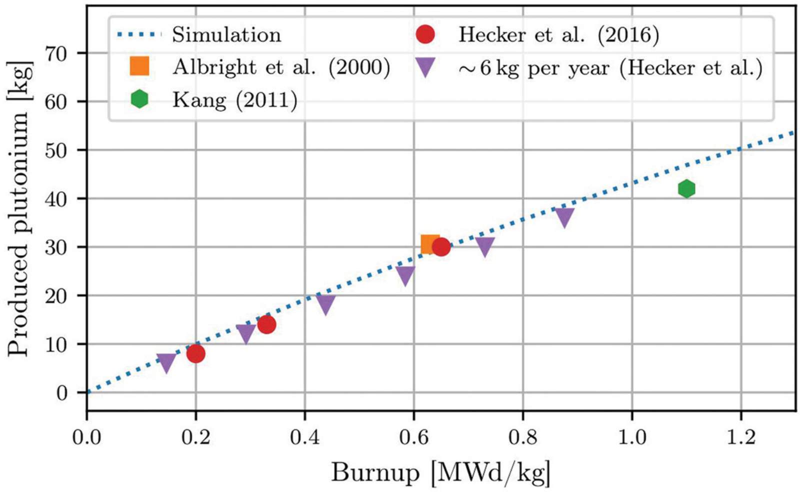

Figure 6. The plutonium produced by the simulated model is displayed dependent on burnup. Plutonium density values of one fuel element are upscaled to the total reactor core. For comparison assessments of various publications are shown.

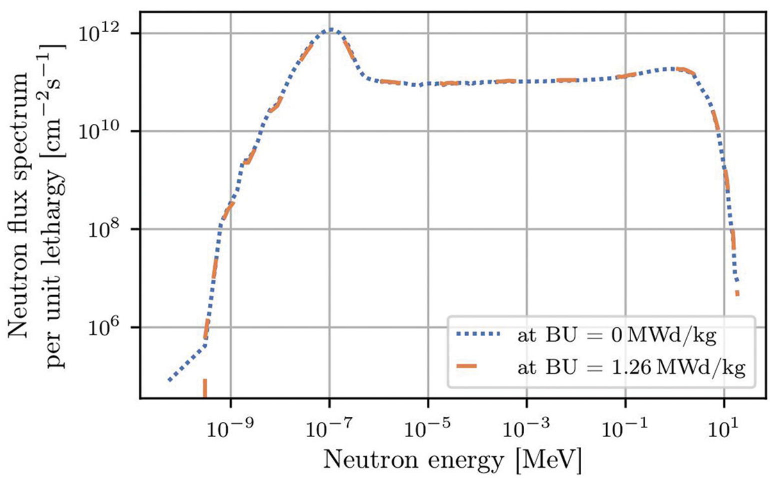

Figure 7. Neutron flux spectrum per unit lethargy at the first and last burnup step in a simulation of the Yongbyon reactor model. The difference between the two spectra is small enough to justify that the spectrum is approximated by averaging over the spectra of each burnup step.

To justify using a time-averaged neutron flux, the neutron spectra at the beginning and at the end of one irradiation cycle were compared. As Figures 7 and 8 show, the change in the flux spectra is small throughout one irradiation cycle.

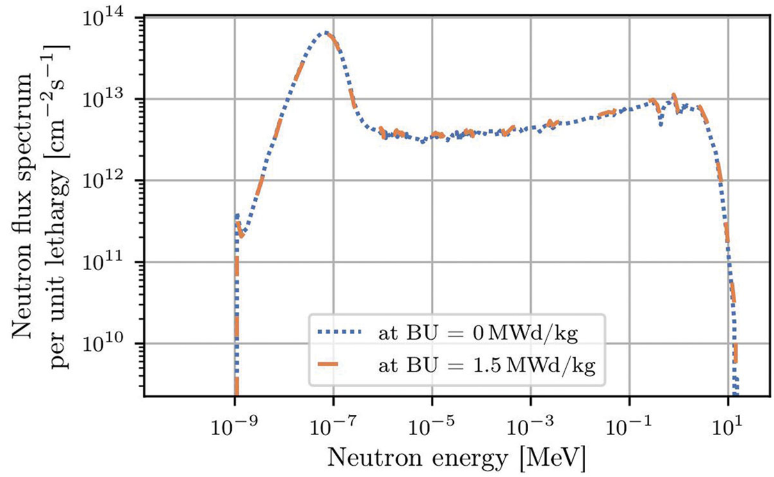

Figure 8. Neutron flux spectrum per unit lethargy at the first and last burnup step in a simulation of the CANDU reactor model. The difference between the two spectra is small enough to justify that the spectrum is approximated by averaging over the spectra of each burnup step.