Statistical Approaches for Pebble Bed Reactor Operations and Safeguards

Philip Gibbs, Oak Ridge National Laboratory
Cory Ball, Oak Ridge National Laboratory

Abstract
The design of pebble bed reactors (PBRs) and their method of operation align more closely with statistical approaches used in manufacturing and process control than traditional safeguards statistical approaches. The reason is PBRs will employ a nondestructive assay (NDA) measurement (burnup measurement system [BUMS]) that is part of the fuel handling system supporting discharge decisions in addition to the reactor code which monitor performance. The integration of these two approaches provides the opportunity to monitor reactor performance statistically for both operations and safeguards in ways not achievable using other reactor designs.

For light water reactors (LWRs), knowledge about reactor code performance in predicting irradiated special nuclear material (SNM) content historically was only achieved from special measurement campaigns or from fuel reprocessing. Conversely, through statistical comparison of the BUMS with the reactor code-predicted values, PBRs can achieve this in real time. The resulting SNM distribution is an indicator of reactor performance because factors such as transit time and path of the pebble fuel through the reactor determine the plutonium production and uranium depletion. By analyzing the predicted and measured values, opportunities exist to adjust operating parameters, fuel design, and other characteristics to optimize performance and fuel utilization. From a safeguards perspective, this approach also provides the information necessary to validate declared values and evaluate whether the reactor is being operated as expected. This paper outlines statistical approaches for PBRs that can be used to support both operations and safeguards.

Statistical Approaches for Burnup Measurements and Calculations
This paper discusses statistical approaches for the measurement systems (i.e., BUMS) and reactor codes that are expected to be used in PBRs. At this point, the relationship between the reactor codes and BUMS with respect to material control and accountability (MC&A) declared values is still being evaluated. Specifically, it has not been determined if (A) BUMS will be used to both determine the burnup that is used to decide when to discharge a pebble and provide the values used in the SNM declarations for burnup/production or (B) if BUMS will strictly support the burnup discharge decision, and the reactor codes provide the values for the SNM declared values.

Either approach may be adequate for the purposes of MC&A because of the diluteness of the SNM in the fuel and the fact that MC&A's main goal is detection of theft or diversion that is based on the removal of a whole pebble. The differences in total plutonium in a spent fuel pebble is so small (i.e., ~0.25 g) that slight variations in this amount have no practical impact on how the MC&A and security system are designed to accomplish that goal.

As a reference, LWRs provide a declared value of SNM in the irradiated nuclear fuel based solely on reactor codes. The regulatory guidance for this practice is provided in ANSI N15.8-2009 [1] Section 9 on SNM calculations for power reactors.

> **9 SNM Calculations**
> **9.1 Element and Isotopic Computations**
> Methods of computation shall be established and utilized for determining the total element and isotopic composition of SNM in irradiated nuclear fuel assemblies and fuel components. The computed values are the basis for shipment documents, as required in 10 CFR 74.15, and material status reports, as required in 10 CFR 74.13.
>
> **9.2 Analysis of Results**
> Refinement of the element and isotopic computations used in determining the SNM content of irradiated fuel should be considered as new technologies evolve. For reprocessed fuel, this may include a collection and comparison of reprocessing plant measurement data with computed data for fuel assemblies.

However, PBRs will employ an NDA measurement (BUMS) that is part of the fuel handling system, which can be used in addition to the reactor code. Therefore, there are two possible "methods of computation" that could be used in determining the SNM content in the irradiated pebbles. One approach would be to use reactor codes like LWRs. The other approach would be to use the NDA or BUMS measurement. It remains to be determined which method, or combination of the two, will be the most accurate predictor of the SNM content in the irradiated pebbles.

Additionally, integration of these two approaches provides a unique opportunity to monitor reactor performance. For LWRs, knowledge about reactor code performance in predicting irradiated SNM content historically was only achieved from special measurement campaigns or from fuel reprocessing (see Section 9.2 ANSI N15.8-2009). Conversely, through statistical comparison of BUMS with the reactor code-predicted values, PBRs can achieve this in real time. The resulting SNM distribution is also an indicator of reactor performance because factors such as transit time and path of the pebble through the reactor determine the plutonium distribution and uranium depletion. By analyzing the predicted and measured values, this information can provide opportunities to adjust operating parameters, fuel design, and other characteristics to optimize performance and fuel utilization.

From a process control and MC&A perspective, there are three statistical decisions to be considered for PBRs.

1. Burnup measurement discharge decision
2. Burnup measurement versus reactor code comparison
3. Analysis of Variance between reactors in a modular multi-unit deployment

1. Statistical Model for Burnup Discharge Decision

This section focuses on the burnup measurement uncertainty associated with the decision on when to discharge a pebble. PBRs have a maximum allowed burnup for the fuel plus a burnup threshold that serve as the decision point for discharge.

One goal of the BUMS is to determine if the current burnup has reached the discharge threshold. The relationship between the discharge threshold and maximum burnup is a buffer to make sure that a pebble is not returned to the reactor such that if it took the highest energy path it might exceed the maximum allowable burnup.

The applicable uncertainties are for both the burnup measurement system and the reactor models, which provide an estimate of the highest energy path taken by a pebble. The remainder of this section will focus on the burnup measurement and how to set the decision point for discharge.

The decision point for discharge is a balance between Type I and Type II[1] statistical errors:

- Type I error—Discharging a pebble when it should have been returned to the reactor, resulting in underutilized fuel.
- Type II error—Returning a pebble to the reactor when it should have been discharged, resulting in a pebble exceeding the maximum desired burnup and creating possible safety concerns and/or less-than-desirable operational performance.

In some situations, such as burial of nuclear waste, the statistical approach used is the measured value plus the measurement uncertainty at 2σ must be less than the maximum allowed limit. Using this approach, a decision threshold would factor in the maximum allowed (or never exceed burnup) minus the maximum likely burnup from an additional pass if placed back in the reactor. The measured burnup plus 2σ would then be less than or equal to that limit.

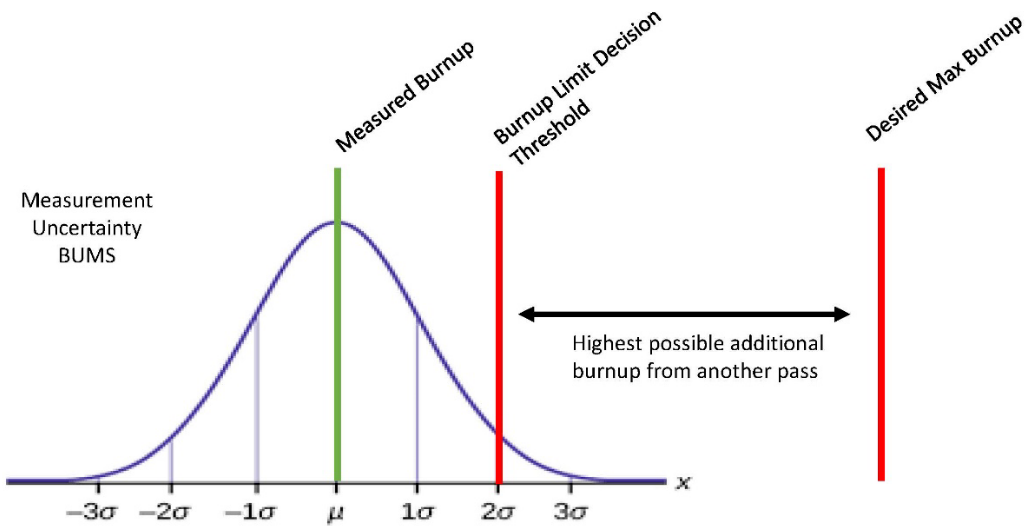

Figure 1. Two sigma limit illustration.

The issue with the 2σ limit approach is that statistically there will be pebbles that will inadvertently be placed back in the reactor that should have been discharged. (This would be a Type II error, assuming it was below the target threshold when it really is not.)

[1] A Type I error (false positive) occurs if an investigator rejects a null hypothesis that is actually true in the population; a Type II error (false negative) occurs if the investigator fails to reject a null hypothesis that is actually false in the population.

In some processes, such as the 6σ approach in quality or other approaches used in criticality, an additional buffer is added to allow for the measurement method uncertainty and reduce the probability of Type II errors. By setting the rejection threshold at a 4σ to even 6σ buffer, the probability of a Type II error can be minimized to whatever extent is desirable. However, this is at the expense of potentially underutilizing the fuel.

Figure 2 provides an illustration of what this would look like statistically. Although the burnup limit detection threshold and desired max burnup remain the same, the average measured burnup target is reduced by additional standard deviations to create a larger buffer reducing the probability of Type II statistical errors.

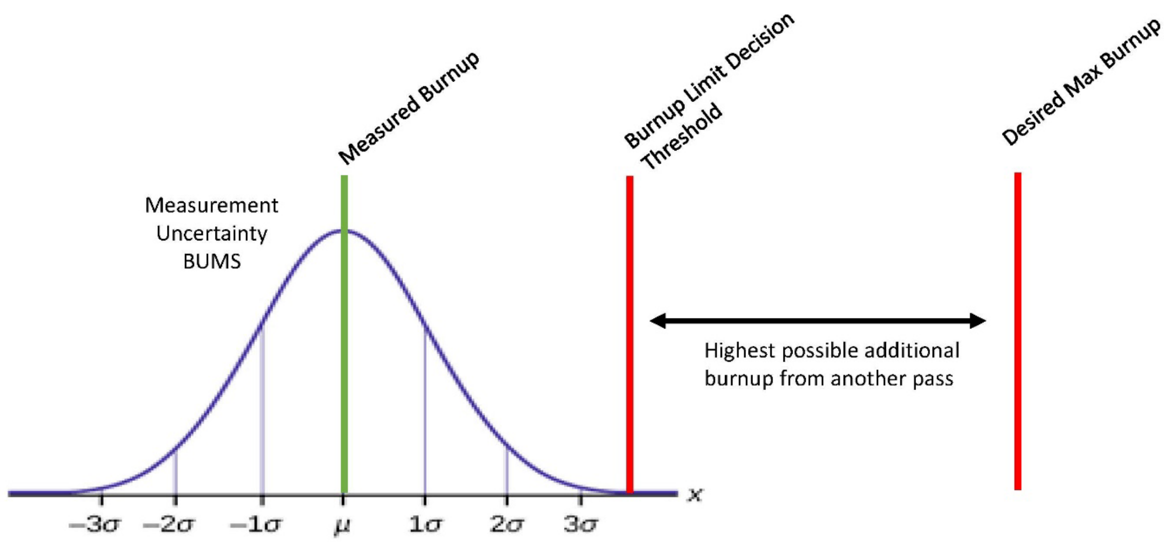

Figure 2. Setting limit to reduce Type II errors.

In summary, the burnup discharge decision is a function of the BUMS measurement uncertainty, the potential maximum energy path through the reactor, and the maximum allowable burnup. The decision point for measured burnup is set by factoring in a margin for the BUMS measurement uncertainty. The size of that margin is a balance between Type I and Type II statistical errors.

2. Burnup Measurement versus Reactor Code Comparison

This section discusses the statistics for comparison and integration of the BUMS measurements for SNM content in the spent fuel with the reactor code predicted values as discussed in the previous section. The MC&A and process benefit of this approach allows for comparison of these two independent methods of SNM content determination from which information can be obtained to support both US domestic and international safeguards, disposition of spent fuel, and reactor performance/operations. It also provides a pathway for continuous process improvement of key aspects of this technology.

The goal is to statistically compare the predicted pebble SNM content from the reactor code with the measured SNM content, to identify any significant differences between the two, and to investigate and identify the underlying cause(s) of the differences. Once the cause is identified, changes can be implemented as needed to adjust the process and eliminate the cause. In some cases, this could be updates to the reactor code and/or underlying assumptions about the pebble flow paths in the reactor. Adjustments to the underlying assumptions supporting the BUMS calibration could also be made.

For illustration purposes using a normal distribution, Figure 3 shows the differences in the statistical distributions that could be expected from these two approaches. Although the distributions overlap, they are statistically different from each other.

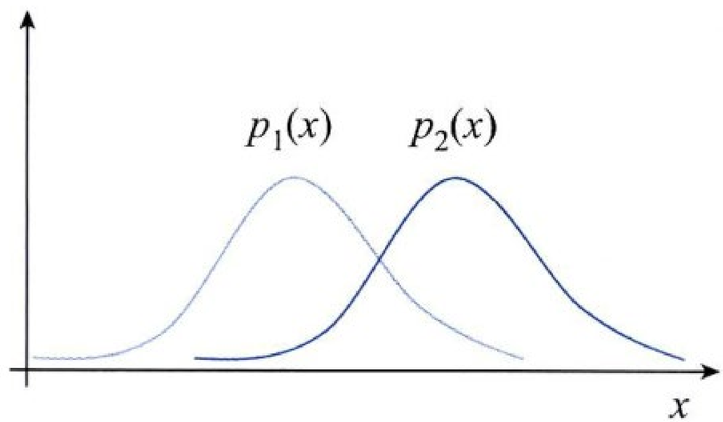

Figure 3. Two different but overlapping distributions.

The desired end state from this approach is shown in Figure 4, where the two approaches for spent fuel SNM content produce essentially the same results with only minor systematic differences.

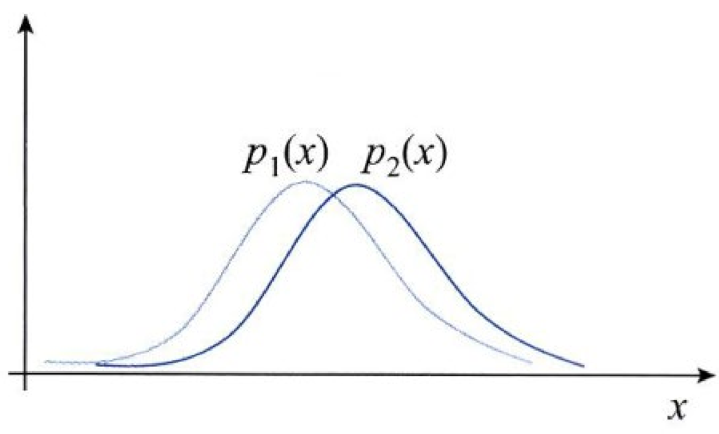

Figure 4. Similar overlapping distributions.

Figure 5. Continuous improvement to reduce process variability.

The goal of this approach is to implement incremental process improvements as needed to reduce process variability to desired levels as determined by key parties (Figure 5). Lastly, this approach also provides independent or redundant quality checks on the process to monitor performance. Through this statistical process control approach, changes in the process can be identified and corrective actions can be taken as needed.

2.1 Statistical sampling plan

First, implementation of this approach could simply be based on comparison of the BUMS measurement with the reactor code for a 100% comparison without a sampling plan. However, because the BUMS measurement is not currently a direct measurement of the SNM content, it may be desirable, at least initially, to sample pebbles and perform more extensive nondestructive or destructive analysis, which would serve to validate the calculated and measured burnup and other values. Typically, these additional measurements are more time-consuming or expensive and therefore the pebbles selected for analysis is a subset of the pebble population.

The International Organization for Standardization (ISO) offers two series of standards for sampling schemes based on acceptance quality limits: (1) ISO 2859 Sampling Procedures for Inspection by Attributes and (2) ISO 3951 Sampling Procedures for Inspection by Variables. In these standards, the acceptance quality limit is defined as the “quality level that is the worst tolerable process average when a continuing series of lots is submitted for acceptance sampling.” Both standards are designed to “ensure that lots of acceptable quality have a high probability of acceptance and that the probability of not accepting inferior lots is as high as practicable.” To this end, the standards suggest when to switch between normal inspection, tightened inspection, and reduced inspection [2, 3].

The major difference between the two series of standards is how a sample is determined to be acceptable. In ISO 2859, acceptance sampling is by attribute: “inspection whereby either the item is classified simply as conforming or nonconforming with respect to a specified requirement or set of specified requirements, or the number of nonconformities in the item is counted." Whereas, in ISO 3951, acceptance sampling is by variables: "acceptance sampling inspection in which the acceptability of the process is determined statistically from measurements on specified quality characteristics of each item in a sample from a lot." Furthermore, ISO 3951 requires the measured variable to be distributed according to a normal distribution or a small deviation from normal [2, 3].

2.2 Implementation

In this analysis, we assume that the sampling plan will be used to select pebbles that will be measured to detect ^{137}Cs activity, which is correlated to burnup values. In this case, ISO 2859 would not be applicable because it is based on acceptance by attribute. Consequently, we should consider ISO 3951, which is based on acceptance sampling by variable. It requires the measured variable to be distributed according to a normal or near normal distribution. However, the results from pebble burnup simulation indicate that the ^{137}Cs activity levels from a random sample of pebbles is unlikely to follow a normal distribution. The results of work done by Oak Ridge National Laboratory on the pebble burnup simulation showed a linear estimation of burnup with ^{137}Cs (Figure 6), which can be used to create the plot shown in Figure 7. The distribution of ^{137}Cs activity levels in the plot clearly does not follow a normal distribution (bell curve). Hence, the ^{137}Cs activity levels from a random sample of pebbles is unlikely to follow a normal distribution.

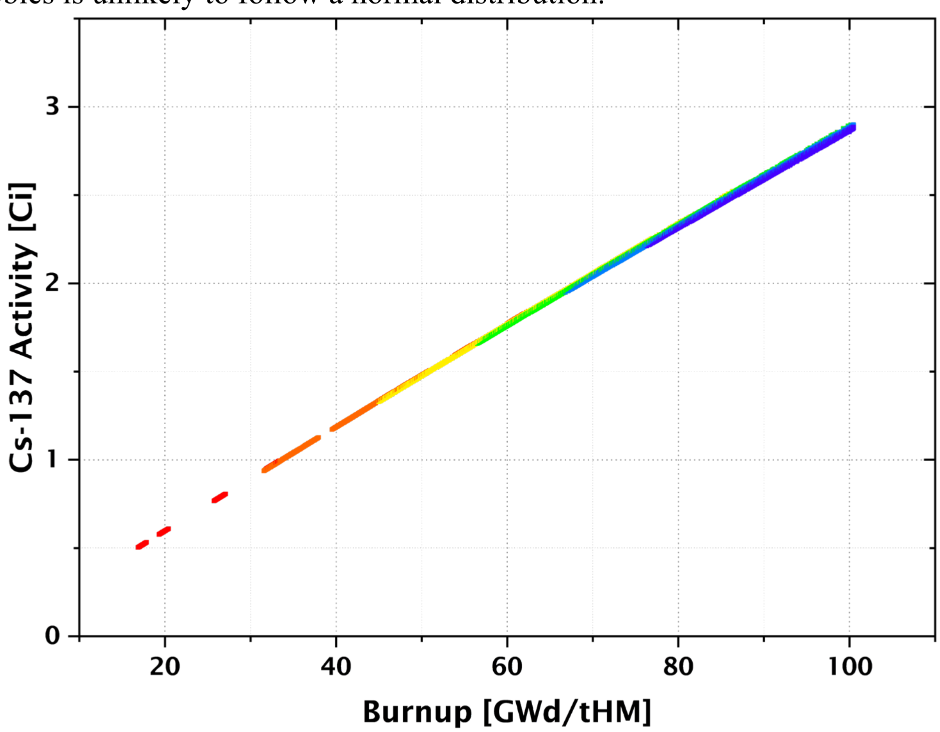

Figure 5. ^{137}Cs activity as function of burnup.

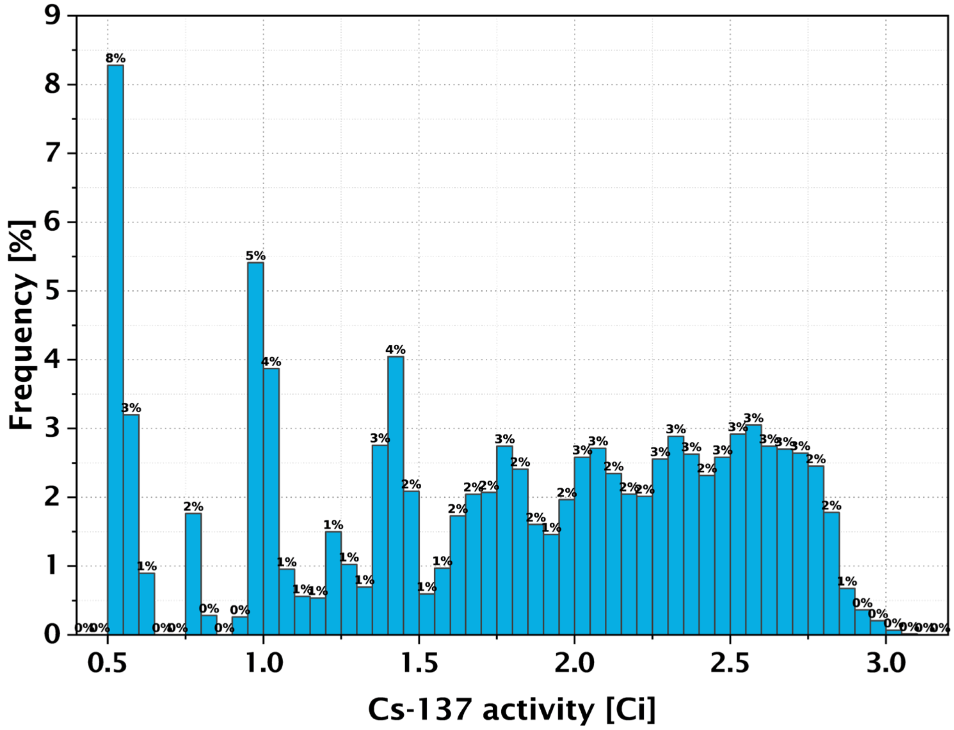

Figure 7. Distribution of measurement for ^{137}Cs.

One approach to consider is taking random sets of samples, computing the means of those sets, and treating those means as the individuals. With a sufficiently large sample, the central limit theorem should apply, and the mean would follow a normal distribution. However, there are two issues with this approach. First, in this case, it is not clear that the mean of the ^{137}Cs activity is necessarily what we need to monitor as opposed to the distribution of the samples. Second, ISO 3951 is based heavily on Shewhart-type control charts. Studies, such as by Huberts et al. (2018) show that these control charts are sensitive to the normality condition [4]. So, it is not advisable to implement ISO 3951 as is.

A solution is to modify the recommended approach in ISO 3951 in such a way that a distribution-free control chart could be used in place of the charts that assume normality. Distribution-free control charts can be used effectively when the underlying distribution of a process is unknown or complex. One such scheme is the location variable and scale chart introduced in [5].

2.3 Code Data Validation

Goodness of Fit Tests

A nonparametric goodness of fit test must be performed to determine if the ^{137}Cs activity samples taken with BUMS are from the same unspecified distribution as those predicted by the model. The null hypothesis of these tests is that the data comes from the same distribution. Type I error is when the null hypothesis is wrongly rejected. To prevent Type I errors, the test statistic is compared to a critical value that depends on sample size and a significance level α, where α is the probability of Type I error.

Kolmogorov–Smirnov
The first test for determining whether two sets of random samples come from the same distribution is the two-sample Kolmogorov–Smirnov test. This test compares the empirical distributions at each point and uses the maximum difference as the test statistic. Given samples of size n, m, respectively, the formulation of the statistic is simple: $D_{m,n} = sup_x\left|F_{1,n}(x) - F_{2,m}(x)\right|$, where $F_{1,n}(x)$ and $F_{2,m}(x)$ are the empirical distributions for the first and second samples, respectively.

The null hypothesis is rejected if

$$D_{m,n} > \sqrt{-\ln\left(\frac{\alpha}{2}\right)\frac{1+\frac{m}{n}}{2m}}.$$

Anderson–Darling
A second test for determining whether two sets of random samples come from the same distribution is the two-sample Anderson–Darling test. The test statistic for this test is more complicated than the Kolmogorov–Smirnov test. For samples of size n and m with $n + m = N$, the test statistic for the two-sample Anderson–Darling test is

$$A^2_{n,m} = \frac{nm}{N}\int_{-\infty}^{\infty} \frac{\{F_n(x) - G_m(x)\}^2}{H_N(x)\{1 - H_N(x)\}} dH_N(x),$$

where $F_n(x)$, $G_n(x)$, and $H_n(x)$ are the empirical distribution functions of the first, second, and combined samples, respectively. Tables for various sample sizes are available in [6].

The two-sample Anderson–Darling test is a more powerful test than the Kolmogorov–Smirnov test. This means given a significance level α, the Anderson–Darling test is less likely than the Kolmogorov–Smirnov test to result in Type II error (i.e., fail to reject a null hypothesis when the null hypothesis is false) [7].

3. Analysis of Variance between Reactor Performance
The discussion so far in this paper has dealt with monitoring the performance of the reactor codes and BUMS for a single reactor. Rather than single reactors, PBRs are modular and are expected to be deployed as multiple units at a given site. It should be expected to see variations in performance between reactor units and systems within those units. Analysis of variance, or ANOVA, is an approach to look at and monitor this variability, make an assessment about its significance, and initiate corrective action if necessary.

Implementing ANOVA would demonstrate that PBRs can dependably attain the same level of reliable performance reactor to reactor. Which ultimately supports widespread deployment of PBRs to address long-term energy and climate goals.

ANOVA is an exceptionally large topic and would be difficult to cover in full detail in this paper. Jaech (1973) provides additional information and a few simple examples that are likely applicable [8]. There are also other publications that can be consulted, and this project will be collaborating with vendors and academia to further explore application of ANOVA to a multi-reactor deployment in FY 2023.

Acknowledgment
This work was carried out in the framework of the Pebble Bed Reactor Nuclear Material Accounting program supported by the US Department of Energy Office of Nuclear Energy. The authors would like to acknowledge the contributions of the complementary work of Donald Kovacic, Jianwei Hu, Donny Hartanto, William Wieselquist, and Robert McElroy Jr.

References

[1] ANSI N15.8-2009, "Material Control Systems—Special Nuclear Material Control and Accounting Systems for Nuclear Power Plants," February 18, 2009

[2] International Organization for Standardization. (1999). ISO 2859.

[3] International Organization for Standardization. (2013). ISO 3951.

[4] L. C. Huberts, M. Schoonhoven, R. Goedhart, M. D. Diko, & R. J. Does, "The Performance of x-bar Control Charts for Large Non-Normally Distributed Datasets," *Quality Reliability Engineering International*, 979–996, 2018.

[5] A. Mukherjee, P. Qiu, & M. Marozzi, "A Comprehensive Distribution-Free Scheme for Tri-Aspect Surveillance of Complex Processes," *Applied Stochastic Models in Business and Industry*, 1157–1181, 2021.

[6] A. N. Petit, "A Two-Sample Anderson–Darling Rank Statistic," *Biometrika*, 161–168, 1976.

[7] B. M. Boyerinas, "Determining the Statistical Power of the Kolmogorov–Smirnov and Anderson–Darling Goodness-of-Fit Tests via Monte Carlo Simulation," Center of Naval Analyses, 2016.

[8] J. L. Jaech, "Statistical Methods in Nuclear Material Control," US Atomic Energy Commission, Richland, WA, 1973.