Disclaimer:
Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.
Heat Source Plutonium: A Case Study in Applying Whisper

Natasha Glazener, Alan Yamanaka, Ray Sartor, James Kuropatwinski

Los Alamos National Laboratory

nglazener@lanl.gov

Introduction:

This paper presents potential points of confusion for the Nuclear Criticality Safety Community to consider and discuss as Sensitivity/Uncertainty analysis codes such as Tsunami and Whisper enter the community’s toolbox. These lessons were learned during a recent evaluation of Heat Source Plutonium operations. First, background of the Heat Source Plutonium operations and the associated criticality safety concerns will be covered. Next, three individual issues will be detailed: what will constitute sufficient additional comparisons as the community grows accustomed to Sensitivity/Uncertainty analysis, how the community would like to see the data, particularly the negative results, communicated, and lastly, a discussion on the difference between understanding the mathematics around a number and understanding its meaning.

Background:

The processes under evaluation were the purification of heat source plutonium, the processing of the plutonium into Radiological Thermoelectric Generators (RTGs), and various characterization and inspections of the final product. For the purpose of these activities, heat source plutonium is defined as having a $^{238}\text{Pu}$ content of at least 65% with the remainder assumed to be $^{239}\text{Pu}$. $^{238}\text{Pu}$ is known to have greater absorption than fission in the thermal region. As a result, $^{238}\text{Pu}$'s minimum critical mass is a single metal piece. Getting a critical metal system is theoretically possible, but due to the high radiation dose and the thermal heat generation of ~0.5 W/g, this has historically not been considered a practical nuclear criticality safety concern. The resulting tribal knowledge was that due to the high concentrations of $^{238}\text{Pu}$, the radiation protection concerns and heat generation concerns bound any potential criticality safety concern for the heat source plutonium.

As greater detail has been added to the regulatory structure, the challenge has become justifying this understanding in a compliant manner. While there have been a variety of experiments elucidating the $^{238}\text{Pu}$ cross sections, the integral experiment with the greatest concentration of $^{238}\text{Pu}$ was ~10 g of $^{238}\text{Pu}$ in a ~5 kg dirty Jezebel run. Clearly, this does not really compare to a system that is 65% $^{238}\text{Pu}$. As a result, traditional validation techniques could only justify an Area of Applicability (AOA) including trace quantities of $^{238}\text{Pu}$. The ANSI/ANS 8 series recommend a single parameter subcritical Limits (SPLs) for $^{238}\text{Pu}$ single metal or oxide units, $^{239}\text{Pu}$ single metal or oxide units and $^{239}\text{Pu}$ solutions. Now, $^{239}\text{Pu}$ is bounding of $^{238}\text{Pu}$ under all moderation conditions; therefore, one strategy is to model everything as $^{239}\text{Pu}$. While this is entirely compliant and certainly bounding, the extreme conservatism of this approach in the solution models obscures where the actual safety margin. Worse, some among the operations personnel had come to regard the limits derived from this methodology as being “compliance” rather than “safety” limits to the
point that they did not understand that there was a nuclear criticality “safety” concern in heat source plutonium. These operators remembered the DantSys curves generated in 2002 (See Figure 1 for MCNP6 reproductions) before the current standard for code validation was written; but the communication of the information had become garbled. In order to get a more transparent limit set, a way to compliantly credit the $^{238}\text{Pu}$ in the solutions was needed.

![Figure 1 Minimum critical mass as a function of concentration for various ratios of $^{238}\text{Pu}$ to $^{239}\text{Pu}$](image)

Whisper demonstrated that a solution containing a mixture of 65% $^{238}\text{Pu}$ and 35% $^{239}\text{Pu}$ has very similar neutronic properties to some of the poisoned $^{239}\text{Pu}$ solution benchmark experiments allowing the evaluation team to quantify and justify a single parameter subcritical limit for heat source plutonium in solution of 6.35 kg rather than the $^{239}\text{Pu}$ solution SPL of 450 g.

**Lesson 1 : How Many Comparisons are Useful?**

While eventually the Sensitivity/Uncertainty codes should be useable with due care but pre-done verifications, the fact that people’s lives rest on this information being bounding likely translates into a long transition period. During this time, it is likely that additional verifications will be requested as people grow accustomed to using this type of validation technique. While having multiple, reinforcing analyses definitely increases the confidence in the results, beyond a certain point, it also increases confusion.

Unsurprisingly, every analyst involved in this endeavor had a different idea of what comparisons to make to verify Whisper was performing as expected. It is important to note that this heat source plutonium analysis was performed with a Version 1.0.0 of Whisper. Whisper version 1.1 will be available with the next MCNP release. As a result, extra caution
was being exercised. When the extra caution of using this early version of the code was combined with the caution for using a new tool, the analysis ended up including discussion of:

- the experimental basis of the $^{238}$Pu cross sections
- comparison of the $^{238}$Pu cross sections to the $^{239}$Pu cross sections
- comparison of the calculated critical masses to known and accepted values for critical and subcritical masses
- comparison of $k_{\text{inf}}$ values calculated in MCNP against the $k_{\text{inf}}$ values calculated using hand calculational formulas on the NNDC data
- comparisons between the macroscopic cross sections of the application models and the benchmark models
- comparisons between the sensitivities of the application models and the benchmark models.

Readers of the document found this plethora of comparisons and data confusing.

While some of these comparisons will become superfluous with the completion of end-user verifications as necessary for local software quality assurance procedures, the community as a whole should discuss what will constitute sufficient verification as the community grows accustom to this new technique.

**Lesson 2: Discussing Negative Results**

The statistical underpinning in the Sensitivity/Uncertainty analysis codes is valid only if the benchmark cases have a sufficient degree of similarity to application models. The variable used to quantify the degree of similarity is correlation coefficient ($c_k$). A $c_k$ of 1 is identical and 0 is no similarity. While numbers can be crunched and values can be produced outside for dissimilar models, the results are not considered meaningful. There is still some discussion of exactly where to draw the line, but the general range is that the $c_k$ should be between 0.8 and 0.9 at a minimum. The evaluation team chose a cut off of $c_k$ values for all of the selected benchmarks >0.9. For application models with a major isotope that effectively does not appear in the benchmark experiments, $c_k$s that do not meet this criterion are to be expected. Therefore, part of verifying that Whisper was performing as expected included a discussion of getting low $c_k$ values in application models that have greater concentration of $^{238}$Pu and lower thermalization of the neutron population.

How to phrase this discussion became a very hotly contested issue. The general pattern that emerged was that a reader’s initial response was to focus on the fact that for the majority of the concentrations investigated, the $c_k$ criteria was not met and that this data should not be used for safety analysis. It generally took a verbal discussion to communicate that this data was being used not for the safety case, but to verify the code was functioning as anticipated. Once the concept was successfully communicated verbally, every analyst involved had different, and frequently mutually contradicting, ideas on how to better communicate this information in the written document. One of the most contested points in
this discussion was the appropriate words to use. An example would be what to call the areas that Whisper is not expected to produce valid results. These are not areas or cases where Whisper “fails”, as Whisper did exactly what it was supposed to do. Nor did Whisper produce “unusable” or “useless” results, because the results were very useful in demonstrated that Whisper did exactly what was expected of it.

Again, having the discussion of how we as a community would like to communicate this information will allow for clear communication of the safety story in future documents.

**Lesson 3: Understanding How a Number was Derived is not the Same as Understanding What it Means**

Thanks to the tutelage of the Whisper creators, by the time NCS got access to Whisper, there was a strong foundation of the mathematics underpinning the code. But when the evaluation team started to apply Whisper to the heat source plutonium evaluation, it quickly became apparent that there was a lack of understanding of what the various variables represent. This was not really an issue in the initial running of Whisper and calculating an SPL as these variables are common in criticality safety. But as the single case end-user verification process forced LANL NCS analysts deeper into the code and results, it became apparent that while we had a basic understanding of the mathematics behind the numerical value, we were unsure of the meaning behind the numbers or how we could use that information. The best example of this is the sensitivity data itself. The name sensitivity data tells you that these numbers tell you which isotopes will have the biggest impact on the system; but what exactly does it mean to have a sensitivity of $10^{-3}$? Does this even have a meaning outside of a comparison? How large of a difference in sensitivities is significant? In the end, we used simple raw comparisons of isotopes in various models. A greater understanding might allow for a more nuanced analysis at future dates.

**Conclusion:**

Hopefully, these points can start the discussions that will allow us to utilize these tools to the fullest.