

Copyright
by
Brian Allen Collins
2022

**The Dissertation Committee for Brian Allen Collins Certifies that this is the
approved version of the following Dissertation:**

**A Statistical Method for Attributing Plutonium Samples to a Reactor
Type from Isotopic Data**

Committee:

Derek Haas, Supervisor

Michaele Brady

William Charlton

Sheldon Landsberger

**A Statistical Method for Attributing Plutonium Samples to a Reactor
Type from Isotopic Data**

by

Brian Allen Collins

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2022

Dedication

This dissertation is dedicated to everyone who has been a constant source of support and encouragement during this process. A special thank you to Ayla, Eilam, and Isaac for giving me nights and weekends to finally complete this.

Acknowledgements

I would like to acknowledge the support of my advisor Derek Haas during the process of developing this dissertation. Being a distance student almost 2,000 miles away has its challenges and Derek helped me navigate those.

I would also like to acknowledge my co-supervisor Mikey Brady for ultimately keeping me on track and providing support and encouragement while we advanced through this work.

Additionally, this work could not have been completed without the help of some amazing statisticians. Ryan Hafen and Landon Segó were instrumental in the early phases of research and analysis by helping me organize the large amounts of data into more usable forms and determining if reactor classification was even possible. Pat Heasler was extremely patient with me and my questions in data analysis and regression. His background in statistics with a nuclear flavor was an excellent resource and without his knowledge and support, applying these methods to this unique problem set would have been extremely difficult.

Finally, I want to thank Chris Gesh and Alex Misner for voluntarily reading this and providing constructive feedback, and Carrie Varley and Sandy Thompson for some final recommendations.

Abstract

A Statistical Method for Attributing Plutonium Samples to a Reactor Type from Isotopic Data

Brian Allen Collins,

The University of Texas at Austin, 2022

Supervisor: Derek Haas

Attributing plutonium to a specific country or organization is a complex and challenging problem that has great interest in the nuclear forensics and counterproliferation communities. Since plutonium is made in a nuclear reactor, identifying the type of reactor, material age, and other physical or chemical characteristics can help in identifying the material origin. For precise attribution, samples or calculations would be needed that represent all operating conditions, for all reactor types, and all possible fuel variations to create a database to be queried to identify a reactor of origin from sampled material. A database this complex is currently unachievable so existing material databases and validated models must be leveraged with new approaches for identification. Since the produced plutonium isotopics are a function of multiple reactor operating parameters (fuel type, fuel enrichment, moderator, local fuel and moderator temperatures, reactor power, and irradiation time), a multi-variate approach is necessary to capture the variation.

In this work, a novel classification algorithm based on regression models of measured and calculated plutonium isotopic data has been developed. While regression analysis is an established method, this is the first application of this technique to available

used fuel plutonium isotopic measurement data combined with calculated data from reactor physics models. The innovative algorithm can quickly identify the most probable reactor type of origin and can ultimately help focus limited resources in the event attribution of interdicted plutonium is necessary.

Measured used fuel isotopic data was obtained through the Spent Fuel COMPOsition (SFCOMPO) database and combined with additional plutonium measurements from the Hanford plutonium production reactors to create a catalogue of plutonium isotopics. The augmented dataset includes measurements and uncertainties where available of the plutonium isotopes ^{238}Pu , ^{239}Pu , ^{240}Pu , ^{241}Pu , and ^{242}Pu for light water moderated reactors, graphite moderated reactors, and heavy water moderated reactors to develop the classification algorithm for use in discriminating the reactor of origin. The developed algorithm can be used to triage plutonium isotopic information to differentiate materials originating in reactors with different moderators, and the potential to discriminate between reactor types with the same moderator.

The new capability provided by the classification algorithm can be applied to real-world scenarios in the nuclear forensic, counterterrorism, and counterproliferation communities. This method can incorporate additional datasets to increase the accuracy of identification as well as expanding the number of different reactor types.

Table of Contents

List of Tables	xiii
List of Figures	xvi
INTRODUCTION.....	1
Problem Statement	1
Background.....	5
Analysis of Measured Data.....	8
Spent Nuclear Fuel Assay Data for Isotopic Validation (NEA, 2011)	9
Plutonium Signatures (Luksic A., 2012).....	9
SFCOMPO Database (Michel-Sendis F., 2014).....	10
Development of Technical Nuclear Forensics for Spent Research Reactor Fuel (Sternat M., 2012)	10
Experimental and Computational Assessment of Trace Nuclide Ratios in Weapons Grade Plutonium for Nuclear Forensics Analysis (Swinney M., 2015).....	11
Operator declaration verification technique for spent fuel at reprocessing facilities (Charlton W., 2000)	12
Modeling and Data Analysis.....	13
Reactor Physics Modeling (Collins B., 2010)	13
Nuclear Forensics Research by Other Institutions.....	15
Development of Technical Nuclear Forensics for Spent Research Reactor Fuel (Sternat M., 2012)	15
Forensic Analysis of samples from the Nuclear Fuel Cycle (Borg L., 2013)	16
Nuclear Forensics Methodology for Reactor-Type Attribution of Chemically Separated Plutonium (Osborn J., 2018a).....	16

Computational and experimental forensics characterization of weapons-grade plutonium produced in a thermal neutron environment (Osborn J. 2018b).....	17
Operator declaration verification technique for spent fuel at reprocessing facilities (Charlton W., 2000)	17
Assessing confidence in inferring reactor type and fuel burnup: A Markov chain Monte Carlo approach (Burr T., 2005).....	18
Data Analysis.....	19
Graphical Data Analysis	19
Statistical Regression.....	20
Multiple Regression Analysis	22
THEORY	25
Nuclear Interactions and Isotope Production.....	26
Nuclear Interactions.....	27
Scattering Interactions	27
Elastic Scattering Interactions.....	27
Inelastic Scattering Interactions.....	27
Transmutation	29
(n,p) Neutron-Proton Transmutation	29
(n, α) Neutron-Alpha Transmutation	30
Capture Interactions.....	31
Fission.....	31
Nuclear Cross Sections	32
Energy Regions.....	34
Thermal Region	37

Resonance Region.....	38
Fast Region	38
Cross Sections and Reactor Operation.....	39
Reactor Operation	42
Isotope Production	44
Bateman Equations	45
²³⁹ Pu Production.....	45
²⁴⁰ Pu Production.....	46
²⁴¹ Pu Production.....	46
²⁴² Pu Production.....	48
²³⁸ Pu Production.....	48
Fuel and Reactor Conditions Impacting Pu Production.....	50
Uranium Enrichment.....	50
Moderator Type and Temperature	50
Reactor Power Level.....	52
Irradiation time	53
Void Coefficient for Boling Water Reactors	53
CLASSIFICATION MODEL DEVELOPMENT	54
SFCOMPO Data Collection and Plutonium Signatures Analysis	55
SFCOMPO - Collection of Available Measurement Data and Preliminary Characterization	55
Pu Signatures - Generation of Calculated Data to Evaluate Isotopic Trends	63
Pu Signatures – Preliminary Analysis of Calculated Data	64
Pu Signatures – Addition of Uncertainty to the Regression Model.....	65

Pu Signatures – Generating a Regression Model from the Calculated Data	67
Pu Signatures – Testing of Regression Model.....	68
Multivariate Characterization of Measured Data (Version 1)	71
Available Training Data and Regression Model Fit	72
Results of Multivariate Analysis Using Only Measured Data.....	81
Discussion and Future Improvements.....	82
Multivariate Characterization Using Measured and Calculated Data (Interim Version of Classification Tool Development).....	83
Regression Model Updates	84
Regression Fit of Combined Data.....	86
Discussion and Next Steps.....	100
Multivariate Characterization with Consolidated Reactor Types (Version 2).....	103
Multivariate Characterization with Outlier Category (Version 3).....	111
Summary and Development of a Simple Tool for Reactor Type of Origin Classification.....	114
Uncertainty of Sample Input.....	117
Conclusions and Future Work	122
Future Work.....	123
Appendix A.....	125
Appendix B.....	131
Data Input Scripts	131
InputDat.r	131
Input.r.....	134
Data Analysis.....	136

Initial.Analysis.r.....	136
CrossVal2.r	139
Analysis Scripts	147
Fun.r	147
References.....	159

List of Tables

Table 1: The natural abundance of U isotopes and their thermal capture, resonance radiative capture, and fission cross-sections (Chart of the Nuclides, 17th Edition).	39
Table 2: Range of Operating Parameters in Current SFCOMPO/NF by Reactor Type	56
Table 3: Pu Isotopics Generated for Regression Model Test	69
Table 4: Number of Data Points Containing Pu Ratios Compared to Complete Ratio Sets	72
Table 5: Alpha Parameter Estimates from Regression Fit.....	79
Table 6: Covariance Matrix (C_i) Estimates from Regression Fit.....	80
Table 7: Confusion Matrix Developed from SFCOMPO/NF Data	81
Table 8: Average Posteriors Developed from SFCOMPO/NF Data	81
Table 9: Complete Datasets from Measurements and Calculated Data.....	87
Table 10: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types	98
Table 11: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.01$	99
Table 12: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.3$	99
Table 13: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 1$	100
Table 14: Alpha Parameter Estimates from Regression Fit.....	106
Table 15: Covariance Matrix (C_i) Estimates from Regression Fit.....	107

Table 16: Average Posteriors Developed for Six Different Reactor Types with a Consolidated PWR type.....	108
Table 17: Average Posteriors Developed for Six Reactor Types with a Consolidated PWR type for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.01$	108
Table 18: Average Posteriors Developed for Six Reactor Types with a Consolidated PWR type for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.3$	109
Table 19: Average Posteriors Developed for Six Reactor Types with a Consolidated PWR type for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 1$	109
Table 20: Average Posteriors Developed for Six Different Reactor Types and an Outlier Category.....	113
Table 21: Calculation of Posterior Probability from Isotopic Ratios obtained from Fukushima-Daini-2	115
Table 22: Comparison of results for a PWR reactor sample input with different levels of uncertainty applied to the value.....	121
Table 23: Calculation of Posterior Probability for Version 3 of the Classification Algorithm with data from a PWR Type Reactor	125
Table 24: Identification of Y , and transformed Y (Y') along with the transformed Burnup (B')	126
Table 25: Alpha Parameter Estimates from Regression Fit.....	126
Table 26: Covariance Matrix (C_i) Estimates from Regression Fit.....	127
Table 27: Inverse Covariance Matrix (C_i) ⁻¹ Estimates from Regression Fit.....	128
Table 28: Calculation of the Covariance Matrix Determinant, Inverse, and natural log of the Determinant and μ_i and $(Y - \mu_i)$ for the Calculation of the loglikelihood	129

Table 29: Calculation of the $C^{-1}(Y - \mu_i)$ Matrix and the *Quadratic* term for the
Calculation of the loglikelihood.....130

List of Figures

Figure 1: Comparison of $^{241}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ for Different Reactor Types as a Method of Graphical Analysis	20
Figure 2: Elastic scattering interaction with a neutron and a nucleus.....	28
Figure 3: Inelastic scattering interaction with a neutron and a nucleus, with a gamma emitted.....	28
Figure 4: Neutron-proton transmutation interaction: $^{16}\text{O}(n,p)^{16}\text{N}$	30
Figure 5: Neutron-alpha transmutation interaction: $^{10}\text{B}(n,\alpha)^7\text{Li}$	30
Figure 6: Neutron-gamma, radiative capture interaction: $^{239}\text{Pu}(n,\gamma)^{240}\text{Pu}$	31
Figure 7: Example of a neutron induced fission interaction	32
Figure 8: Total, Scattering, Capture, and Fission Neutron Cross Sections in ^{235}U (Chadwick M., 2011)	34
Figure 9: Neutron Cross Sections in ^{235}U with the specific regions identified (Chadwick M., 2011)	37
Figure 10: Fission and Absorption Neutron Cross Sections for Uranium (Chadwick M., 2011).....	40
Figure 11: Fission and Absorption Neutron Cross Sections for Plutonium (Chadwick M., 2011).....	41
Figure 12: Prompt Neutron Energy Spectrum (Lamarsh J., 2001)	42
Figure 14: Moderator Elastic and Absorption Cross Sections (Chadwick M., 2011)	52
Figure 15: $^{238}\text{Pu}/^{239}\text{Pu}$ Plotted for Multiple Reactor Types	58
Figure 16: $^{241}\text{Pu}/^{239}\text{Pu}$ Plotted for Multiple Reactor Types	58
Figure 17: $^{242}\text{Pu}/^{239}\text{Pu}$ Plotted for Multiple Reactor Types	59

Figure 18: PWR, BWR, and RBMK $^{238}\text{Pu}/^{239}\text{Pu}$ Measurement Data Comparison with Second Order Least Squares Fits and Uncertainty Bands	61
Figure 19: PWR, BWR, and RBMK $^{241}\text{Pu}/^{239}\text{Pu}$ Measurement Data Comparison with Second Order Least Squares Fits and Uncertainty Bands	62
Figure 20: PWR, BWR, and RBMK $^{242}\text{Pu}/^{239}\text{Pu}$ Measurement Data Comparison with Second Order Least Squares Fits and Uncertainty Bands	62
Figure 21: Pu Ratios of Calculated Datasets for Commercial Reactors	64
Figure 22: Posterior probability that sample is from tested site given that it came from true site for all four reactor types at a given timestep.....	65
Figure 23: Comparison of $^{238}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ for the Calculated Dataset	66
Figure 24: Posterior probability that sample is from tested site given that it came from true site for all four reactor types with 3% uncertainty at a given timestep	67
Figure 25: Posterior probability that sample is from tested site given that it came from true site for all four reactor types with 15% uncertainty at a given timestep	68
Figure 26: Testing of Regression Model Using Measured Data.....	70
Figure 27: Regression of Different Reactor Types based on $^{238}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ Comparison	76
Figure 28: Regression of Different Reactor Types based on $^{241}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ Comparison	77
Figure 29: Regression of Different Reactor Types based on $^{242}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ Comparison	78
Figure 30: Regression Fit and 2-sigma error bounds of CANDU (Heavy Water) Reactor Type with Measured and Calculated Data.....	89

Figure 31: Regression Fit and 2-sigma error bounds of GE (BWR) Reactor Type with Measured and Calculated Data	90
Figure 32: Regression Fit and 2-sigma error bounds of Hanford (Graphite) Reactor Type with Measured and Calculated Data	91
Figure 33: Regression Fit and 2-sigma error bounds of Magnox (Graphite) Reactor Type with Measured and Calculated Data	92
Figure 34: Regression Fit and 2-sigma error bounds of RBMK-1000 (Graphite) Reactor Type with Measured and Calculated Data.....	93
Figure 35: Regression Fit and 2-sigma error bounds of Westinghouse (PWR) Reactor Type with Measured and Calculated Data	94
Figure 36: Regression Fit and 2-sigma error bounds of VVER-1000 (PWR) Reactor Type with Measured and Calculated Data	95
Figure 37: Regression Fit and 2-sigma error bounds of VVER-440 (PWR) Reactor Type with Measured and Calculated Data	96
Figure 38: Comparison of Regression Fits with Classification Error for the Eight Reactor Types	97
Figure 39: Error rates as a function of Burnup ($^{240}\text{Pu}/^{239}\text{Pu}$) for the different reactor types	102
Figure 40: Regression Fit and 2-sigma error bounds of the Consolidate PWR Reactor Type with Measured and Calculated Data	104
Figure 41: Comparison of Regression Fits with Classification Error for the Six Reactor Types	105
Figure 42: Error rates as a function of Burnup ($^{240}\text{Pu}/^{239}\text{Pu}$) for the different reactor types	110

Figure 43: Screenshot of the Reactor Classification Tool showing Probability of
Reactor Type of Origin based on an Input from a BWR dataset117

INTRODUCTION

Problem Statement

The United States continues to develop nuclear nonproliferation leadership with the intent of reducing the dangers posed by nuclear weapons, as nuclear terrorism is a global threat. (National Security Council, 2021) (DOD, 2018) Plutonium is produced during the operation of nuclear reactors by the irradiation of uranium fuel and can be used both as a nuclear fuel or for weapons purposes. Because of the potential weapons related aspect of the element, resources are spent for plutonium storage, security, and tracking. If plutonium is discovered or encountered outside of regulatory control, it becomes important to identify the source of the material quickly and with relative confidence in the result. Since plutonium is produced in nuclear reactors, being able to identify a reactor type of origin is a practical method for narrowing down the plausible countries or organizations of origin for the material, which has been of great interest in the nuclear forensics and counterproliferation communities. (IAEA, 2015), (IAEA, 2020), (IAEA, 2018), (IAEA, 2017)

The objective of the research in this dissertation is to develop a validated approach to enable identification of the reactor type of origin for plutonium associated with the interdiction of pre-detonation¹ material. In this method, identification relies primarily on measured plutonium isotopic ratios. For the classification algorithm to be successful, sufficient data must be available to develop the methodology, therefore, identifying and including both measured and calculated isotopic data are necessary. The primary focus in this work is on intra-elemental isotopic ratios to mitigate impacts of chemical processing,

¹ “Pre-detonation” refers to nuclear material that is recovered or sampled before nuclear yield has occurred, generally resulting in the ability to directly collect physical, chemical, and isotopic signatures. “Post-detonation” is material that is recovered after nuclear yield is obtained, which results in radioactive samples and debris that need to be analyzed in order to back-calculate the pre-detonation material characteristics.

as that can alter the material characteristics when the plutonium is separated from the used fuel. However, there is still significant information that can be gleaned by analyzing trace element (uranium and fission product) data in potential follow-on work.

The goal is to develop statistical models utilizing experimental data as well as calculated or modeled data to fill in gaps in the measured datasets. These models will be used for the identification of reactor types that could be credible sources of specific nuclear materials obtained from the investigation of interdicted material information or post-detonation material information. The usefulness and reliability of these models will be dependent on the quality and quantity of data upon which they are based. As additional data is obtained or generated, the models can be refined and improved to expand its applicability to other reactor types or operating conditions.

The statistical models that are developed must represent the range of expected normal and known operating conditions for the specific reactor type and provide the user an indication if the plutonium sample is an outlier or falls outside of expected data ranges. Measured used fuel isotopic data available in the open literature is limited, with respect to the range of initial fuel compositions, assembly, and reactor design, etc. Additionally, the data is limited to the specific reactors involved in the data collection. Therefore, not all potential operating scenarios the reactor or fuel type could have been exposed to are represented including potential variations in initial fuel enrichment. Furthermore, fuel measurements are very expensive, require specialized facilities, and are limited in scope. For example, in a pressured water reactor, each fuel pellet (with approximately 10,000,000 fuel pellets in a core) is irradiated under slightly different conditions with these conditions changing over time. There are also issues with modeled or calculated data to make sure they represent the type of reactor they are intended to represent, and that the data is categorized accurately. Capturing all this information is unrealistic from both a modeling

or sampling perspective, therefore additional methods must be developed to utilize the existing data for material characterization and attribution.

Once this method is developed, additional data sources can be leveraged to increase both the range of reactor operations and other reactors or production pathways of interest. These can include measured data or validated model data to fill in gaps in the dataset.

In a perfect situation, there would be physical samples from every individual reactor that represent all known operating conditions, creating a pedigree of material characteristics that could be used to identify a reactor of origin from interdicted material. Since that is unachievable for a variety of reasons, existing material databases must be leveraged, and gaps of information need to be filled with validated models. There are limited databases of used fuel radionuclide inventories suitable for comparison against interdicted materials or for validating and verifying computational models. Using the international Spent Fuel COMPOsition (SFCOMPO) database containing measured data for Pressurized Water Reactors (PWRs) and Boiling Water Reactors (BWRs) as the basis (Kurosawa M., 1997), additional data from the Hanford plutonium production reactors was added to generate a SFCOMPO – Nuclear Forensics (SFCOMPO/NF) database (Hanson B., 2012) (Brady-Raap M., 2011). The SFCOMPO/NF database was further expanded to include measured and simulated isotopic inventories of used fuel from a variety of reactor types including PWRs, BWRs, Canada Deuterium Uranium (CANDU), MAGNOX, Vodo-Vodyanoi Energetichesky Reaktor (VVER) and Reaktor Bolshoy Moshchnosti Kanalnyi, "High Power Channel-type Reactors" (RBMK) (Brady-Raap M., 2012).

The intended use of the data is to investigate the use of ratios of isotopes to discriminate materials derived from these different reactor types. Known attributes of these isotopic ratios for other reactors, such as graphite production reactors, CANDU,

MAGNOX, VVER and RBMK reactors provided expectation that the distinction for these and the group of PWR and BWR fuels would be clear.

The focus of this research is to develop a classification tool for use in preliminary reactor type identification through statistical analysis of both measured and calculated data. This is achieved by comparing plutonium sample data to populated datasets and using “reactor physics models” for the different reactor types to fill in the gaps of isotopic information. These “reactor type models” containing both measured and calculated data will be created for the different reactor types and will be used to develop a “classification algorithm” which can distinguish reactor type of origin based on plutonium isotopic measurements. This classification algorithm is used to develop a tool to rapidly provide the probability of reactor origin for a specified plutonium sample.

Background

There has been a large push in the last decade regarding nuclear forensics research. This research has been mainly identifying the scope of the problem and working towards practical solutions for identification. This is a data intensive problem that unfortunately has more gaps in information than verified and validated data. The National Technical Nuclear Forensics Center (NTNFC) has been funding research in this area for the past decade with annual conferences to share progress in this field. The NTNFC was established on October 1, 2006, and has the following mission:

“The center has three primary missions identified by Presidential Directive and affirmed in the Nuclear Forensics and Attribution Act (Public Law 111-140) which President Obama signed into law on February 16, 2010. These missions are to serve as program integrator and steward for the U.S. Government to ensure a ready, robust, and enduring nuclear forensics capability, to advance capabilities to conduct forensics on nuclear and other radioactive materials, and to lead the National Nuclear Forensics Expertise Development Program.” (DHS, 2022)

The Countering Weapons of Mass Destruction (CWMD) Office was established in December 2017 by consolidating primarily the Domestic Nuclear Detection Office (DNDO) and other Department of Homeland Security (DHS) elements. With NTNFC established under DNDO, its mission space was transferred to the CWMD Office.

As the threat landscape has evolved, the forensic analysis mission space has continued to shift. The Department of Energy (DOE) National Nuclear Security Administration (NNSA) is establishing nuclear forensics capabilities in Defense Nuclear Nonproliferation (NA-20) and Counterterrorism and Counterproliferation (NA-80). The NA-80 mission has also established the Office of Nuclear Forensics to serve as NNSA’s

lead coordinator for Technical Nuclear Forensics capabilities and policy. The following description establishes the Office's capabilities for pre-detonation and post-detonation analysis for the purposes of helping support attribution:

“Nuclear forensic capabilities were first developed for nuclear test monitoring and treaty verification purposes during the Cold War. They are now used for a wide range of national security and law enforcement missions, including determining the origin of nuclear materials outside of regulatory control, such as those seized from nuclear smugglers. If a nuclear threat device were interdicted before detonation, analysis of the nuclear material and device design can yield information about its origin and manufacturer. Should terrorists succeed in detonating a nuclear device, the office has operational capabilities that enable debris analysis and other activities to identify the source of the nuclear material used, as well as provide insights about the device design. Coupled with intelligence and law enforcement information, these capabilities support the process to attribute responsibility for the attack.”
(DOE, 2022)

While the counterterrorism and counterproliferation mission spaces have dominated the arena of nuclear forensics, analysis of used fuel has also been studied for other applications including burnup credit and nuclear criticality safety.

In addition, sensitivity and uncertainty analysis has been identified as a major area in both calculation approaches and computer code development. The identification of parameters of interest and their effect on the solutions generated allow for their use in classification approaches that may be useful in determining origination of material. Multiple approaches have been developed that determine sensitivity and uncertainty and their application will be discussed.

Previous and ongoing work in the nuclear forensics subject area can be categorized under the following categories:

- Interdicted Sample Measurement and Characterization
- Sample Analysis and Attribution
- Post Detonation Examination and Analysis

Without a robust sample measurement and characterization platform, we would not have measured data to validate physics models or apply to analytical methods development. However, that is not the focus of the research and development of the classification tool presented here. For information on sampling methods, characterization techniques, and other measurement processes, please refer to articles such as: “*Nuclear, Chemical, and Physical Characterization of Nuclear Materials*” by (Tandon L., 2008) Tandon, L., Hastings, E., Banar, J. et al.; “*Nuclear Forensics Analysis*” (Moody K., 2015) by Moody, Grant, Hutcheon; and “*Use of Secondary Ion Mass Spectrometry in Nuclear Forensics analysis for the Characterization of Plutonium and Highly Enriched Uranium Particles*” (Betti M., 1999) by Betti, Tamborini, Koch.

The analysis and attribution category includes a large variety of methods to leverage both modeled and measured data with the intent of identifying the origin of the nuclear material. For the purposes of this research, the focus is on identifying specific reactor types of origin as a first step which can be used to focus other efforts in further identifying the source of material.

The following subsections discuss the advancements in nuclear forensics analysis, focusing on:

- Use of measured data as a method for classification or model validation
- Development of models to fill in gaps in data or generating data for classification, potentially incorporating uncertainties in the method

- Use of statistical approaches for classification

This is shown by the work completed in the field of nuclear forensics and material attribution completed for NTNFC since the mid-2000s at Pacific Northwest National Laboratory, combined with specific research from other organizations and researchers for the topics noted above. Additionally, broader research completed by the larger nuclear forensics community in relation to this topic will be discussed.

ANALYSIS OF MEASURED DATA

This section includes the research focused on the utilization of measured data. Many projects leverage commercial data that was available for a variety of uses, of note: The Spent Fuel Composition Database (SFCOMPO) and the State-of-the-Art (SOAR) reports were developed to address nuclear criticality safety in the transportation and storage of spent fuel (Kurosawa M., 1997) (NEA, 2011). How they apply to and have been utilized in this dissertation will be discussed in detail. Additionally, there are multiple areas of research that focus on specific fuel types and material analysis. These can include research reactors, power reactors, and material production reactors. Measured data can be used to develop statistical models for a variety of analyses and validate reactor physics models for the generation of data where measurements do not exist. There are a number of publications that are relevant to the analysis of measured data, including the determination of reactor type from the data points (Mayer M., 2000), (Mayer K., 2005), (Wallenius M., 2000), (Nicolaou G., 2017), (Nikolaou G., 2019), (Hoover A., 2009), (Joshi J., 2019) Some of the more relevant research is discussed in further detail below. Some of this research had

aspects in one or both of the other focus areas and those specific analyses will be discussed in the appropriate section.

Spent Nuclear Fuel Assay Data for Isotopic Validation (NEA, 2011)

This state-of-the-art report sponsored by the Nuclear Energy Agency (NEA) and written by the Nuclear Science Committee, Working Party on Nuclear Criticality Safety, and Expert Group on Assay Data of Spent Nuclear Fuel (EGADSNF) discusses the burnup credit and nuclear criticality safety for used nuclear fuel. The paper discusses the importance of experimental measurements of irradiated nuclear fuel for the purposes of validating computer models. The report discusses the technical activities of the EGADSNF and the associated efforts of many NEA member countries to compile and document a comprehensive database of assay data to support applications related to nuclear energy and irradiated fuel safety and management. This paper discusses the procedures for measured data to be included in the SFCOMPO database for use of computer code development and validation.

Plutonium Signatures (Luksic A., 2012)

Multiple years of research and analysis were completed under NTNFC identifying plutonium signatures. Some of these reports have limited distribution based on the sponsor and area of research. The focus of this effort was to quantify differences in plutonium isotopics based on different reactor operating parameters. Graphite production reactor data was obtained through different sources and included operating data from the Hanford reactors and reprocessing facilities. Combining the data with reactor operation logs, a preliminary assessment on plutonium signatures as a function of operating conditions was

obtained. Additional efforts are underway to find isotopic information for other plutonium production reactors which include the Savannah River Heavy Water Production Reactors.

SFCOMPO Database (Michel-Sendis F., 2014)

SFCOMPO, the Spent Fuel Isotopic Composition Database, was originally developed at the JAERI Department of Fuel Cycle Safety Research's Fuel Cycle Safety Evaluation Laboratory (Kurosawa M., 1997). SFCOMPO provides isotopic composition data via the internet. It archives measured isotopic composition data and the values of their ratios, which are required for the validation of burn-up codes.

Based on discussions at a meeting of the Working Party on Nuclear Criticality Safety (WPNCS) held in December 2001, the initial system of SFCOMPO was transferred from JAERI to the NEA Data Bank for internet dissemination. The current SFCOMPO database contains experimental data coming from 44 different reactors of 8 different international types, currently representing 750 fuel samples. The database currently contains more than 24,000 measurement entries including the following reactor types: Boiling Water Reactors (BWR), Pressurized Water Reactors (PWR), CANDU Heavy Water Power Reactors, VVER Pressurized Water Reactors, RBMK Graphite Power Reactors, and MAGNOX reactors. This database was augmented with isotopic data obtained from the Hanford Declassified Database to include information on plutonium isotopes from the graphite production reactors located at Hanford.

Development of Technical Nuclear Forensics for Spent Research Reactor Fuel (Sternat M., 2012)

This dissertation by Matthew Sternat focused on the development of techniques for the purposes of identifying and reconstructing specific reactor parameters by analyzing reactor used fuel data. Part of this research was comparing modeling and simulation results

with data from a sample that was irradiated and went through radiochemical analysis. For this work, the reactor that was chosen was the Oak Ridge Research Reactor (ORR), which was fueled by 93.1% HEU as U_3O_8 . After locating a suitable fuel assembly, the used fuel was shipped to Savannah River National Laboratory (SRNL) for radiochemical analysis. There were four principal analyses performed on the ORR used fuel samples: ICP-AES, ICP-MS, gamma spectrometry, and alpha spectrometry, with the ICP-MS and gamma spectrometry measurements used in the inverse analysis portion of the dissertation, the specifics of which will be discussed in the Nuclear Forensics Research by Other Institutions section.

This experimental work attempted to account for the uncertainties in the measured data, however the ICP-MS results did not include uncertainties associated with counting statistics. To estimate the uncertainty in the results, an assumption was made for the uncertainty of the ^{235}U signal, as it is the strongest signal in the results and therefore should have the lowest measurement uncertainty. The uncertainty for the other isotopes was then determined by scaling to the ^{235}U results. The measurement data and corresponding uncertainties were applied in the reactor parameter reconstruction techniques of the research.

Experimental and Computational Assessment of Trace Nuclide Ratios in Weapons Grade Plutonium for Nuclear Forensics Analysis (Swinney M., 2015)

This dissertation by Matthew Swinney focused on the development of a methodology for source attribution of interdicted material, specifically plutonium based on fission product measurements. The portion of the work focusing on experimental data involved the irradiation and analysis of depleted uranium dioxide (DUO_2) in the High Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory (ORNL). The samples irradiated

and analyzed were used to validate the reactor physics models, in this case, MCNPX, by comparing the fission product outputs. Specifically, ^{144}Ce , ^{134}Cs , ^{137}Cs , ^{154}Eu , and ^{125}Sb were used as a comparison between the gamma spectroscopic measurements, the mass spectroscopic measurements, and the MCNPX simulation.

This approach is necessary to validate the reactor physics models for isotope production and reactor analysis. If the model is validated by measured data, it can be assumed that the results generated will be accurate across the range of parameters that the model was validated for. This research then applied statistical methods for potential attribution of material. These conclusions will be discussed in the statistical methods section included in the background analysis.

Operator declaration verification technique for spent fuel at reprocessing facilities (Charlton W., 2000)

This journal article focused on a verification technique for use at reprocessing facilities using measured isotopic ratios of stable noble fission gases from on-stack emissions during reprocessing of used fuel using high-precision mass spectrometry. Results from the measurements were then compared to a database of calculated isotopic ratios using a data analysis method to determine specific fuel parameters (e.g., burnup, fuel type, reactor type, etc.). The integrated system (mass spectrometry, reactor modeling, and data analysis) was validated using on-stack measurements during reprocessing of fuel from a US production reactor. This work suggested that the technique of fission gas collection and analysis developed therein may have some difficulty distinguishing pressurized water reactor (PWR) from boiling water reactor (BWR) fuel; however, it consistently can distinguish light water reactor (either PWR or BWR) fuels from other reactor fuel types when deployed at reprocessing facilities.

MODELING AND DATA ANALYSIS

This section focuses on the generation of data through reactor models to fill specific information or measurement gaps. As it is difficult and cost prohibitive to obtain samples of all known used fuel and/or plutonium production operations, modeling provides an opportunity to generate data that can be used to characterize a broader area of expected used fuel. It must be noted that the models used need to be validated for the operating regions that they are being used for and the uncertainty in the results must also be captured to properly quantify the results. Validation of calculated data can be completed by comparing against samples with known operational parameters. This is discussed further in section describing the classification model development using calculated data.

A variety of reactor physics models and codes have been developed and used for this type of analysis. As in the previous section, there are a number of publications that are relevant to this area of research (Wallenius M., 2000), (Kitcher E., 2019), (Hellesen C., 2017), (Dayman K., 2014), (Jones A., 2014), (Charlton W., 2001), with specific research discussed in further detail below. Additionally, various details and methods for using reactor models and simulated data for training models can be found in (Chirayath S., 2009) (Chirayath S., 2015).

Reactor Physics Modeling (Collins B., 2010)

Recognizing that there are gaps in the measured data that is currently available, efforts were made to fill in gaps with calculated data. To determine the effectiveness of computer models and simulation, a project was completed that comprised a parametric study on reactor parameters that were thought to have an impact on final plutonium isotopic composition.

Minor plutonium isotopes may be used to identify the production source (i.e., reactor class) for most worldwide stockpiles of separated plutonium. However, to carry out such classification on an unknown sample would require a comparative database of destructive analyses or accurate model burnup simulations over a range of reactor designs and operating parameters. Given the scarcity of destructive assay measurements of the global plutonium stockpiles, it is imperative to confirm the quality and accuracy of existing reactor codes, which was the focus of this study. The work presented attempted to evaluate the differences between two separate calculational approaches routinely used to predict plutonium isotopics in various reactors: Monte Carlo N-Particle (MCNP) and transport theory (SCALE). Major and minor plutonium isotope inventories predicted from these models were compared as a function of key reactor parameters. This study found that inventories predicted for major isotopes were similar (<5% for ^{239}Pu , ^{240}Pu and ^{241}Pu) for both codes. However, somewhat larger differences between model predictions of the two codes were found (10-20% for ^{238}Pu and ^{242}Pu) for the minor isotopes of plutonium, primarily due to the differences in the cross sections used for the analysis (Monteburns using ORIGEN2.2 data libraries for the MCNP cases and the ENDFv6 cross sections in SCALE5.1). Both models predicted major plutonium isotope inventories relatively accurately. However, the limited availability of measured plutonium isotope inventories in used fuel prevented the authors from identifying a “preferred” burnup model of the two compared in this study for predicting the inventories of the minor isotopes. As more measurements become available, especially those characterizing historic samples, a re-evaluation may be necessary. As discussed in the previous section, validating the output isotopics from either computer code could facilitate the generation of additional calculated data, however, it must be noted that the data generated would only be valid for the operating parameters in which the code was validated against.

NUCLEAR FORENSICS RESEARCH BY OTHER INSTITUTIONS

Multiple research project, journal articles, and dissertations have been completed with the focus of inverse analysis of data and even as far as attempting reactor attribution. A large body of work was completed by Savannah River National Laboratory, Los Alamos National Laboratory, Lawrence Livermore National Laboratory, and Texas A&M University. The following subsections focus on the work of these institutions that have a similar application although different approaches to the work completed for this dissertation.

Development of Technical Nuclear Forensics for Spent Research Reactor Fuel (Sternat M., 2012)

This dissertation by Matthew Sternat focused on the development of techniques for the purposes of identifying and reconstructing specific reactor parameters by analyzing spent fuel. The focus was primarily on research reactors, a reactor type that was not considered for this dissertation, fueled with HEU. Reactor physics simulations were performed to verify the reaction rate and neutron transport behaviors, which were confirmed by the analysis of samples that were irradiated and then measured. Ultimately the goal was to reconstruct the initial fuel composition of the sample and utilized fission products (^{137}Cs and ^{148}Nd) and ^{235}U and ^{238}U for initial enrichment and cooling times. These techniques included the use of reactor physics simulations, analysis of samples, and analytic calculations as methods to either identify or exclude possible sources of nuclear material. Sternat had similar conclusions to other work regarding the accuracy and uncertainties generated with reactor physics models and showed that the generated or measured isotopic information could be used to determine a reactor type, however, the method for back calculating the result is time intensive and requires more information regarding fuel characteristics than the statistical method developed in this work.

Forensic Analysis of samples from the Nuclear Fuel Cycle (Borg L., 2013)

This report summarized different analytical techniques that could be employed in forensics investigations of nuclear materials in all aspects of the nuclear fuel cycle. Starting with the IAEA steps in the nuclear fuel cycle, different techniques were evaluated on the signatures that were created or lost as material moved through the different processes. The research completed may be used to provide additional data of currently held material or specific sample data in the event of interdicted material. The report focused on a variety of methods for collecting physical data on a sample. The different physical techniques that were identified in this technical report could ultimately be used as future discriminators in the classification tool that was developed.

Nuclear Forensics Methodology for Reactor-Type Attribution of Chemically Separated Plutonium (Osborn J., 2018a)

This journal publication by Texas A&M University developed a methodology to attribute separated Weapons Grade plutonium (WG Pu) to a specific reactor type of origin. The development of the method relied on reactor physics models of different reactor types to calculate the isotopics of the used fuel at burnups less than 5 GigaWatt-days/Metric Ton Uranium. This burnup region gives $^{240}\text{Pu}/^{239}\text{Pu}$ ratios of less than 6%. The method focuses on the use of Intra-Element ratios that are unaffected by chemical separation processes. They selected the following isotope ratios to predict reactor parameters: $^{137}\text{Cs}/^{133}\text{Cs}$, $^{154}\text{Eu}/^{153}\text{Eu}$, $^{134}\text{Cs}/^{137}\text{Cs}$, $^{135}\text{Cs}/^{137}\text{Cs}$, $^{136}\text{Ba}/^{138}\text{Ba}$, $^{150}\text{Sm}/^{149}\text{Sm}$, $^{152}\text{Sm}/^{149}\text{Sm}$, $^{240}\text{Pu}/^{239}\text{Pu}$, $^{241}\text{Pu}/^{239}\text{Pu}$, and $^{242}\text{Pu}/^{239}\text{Pu}$. An interesting ratio that is not present is $^{238}\text{Pu}/^{239}\text{Pu}$, as this can be an important distinguisher of the original uranium content in the fuel (if there is ^{236}U present because the fuel was reprocessed). With all the data generated, a maximum likelihood method utilizing a conditional probability density function was developed. It should be noted that using three distinct reactor types (Pressurized Water Reactor,

Pressurized Heavy Water Reactor, and Fast Breeder Reactor) to generate the model data results in a high probability of correctly classifying the reactor type of origin as was noted in preliminary Plutonium Signatures work. Additional data that overlaps may affect the classification capability of this method.

Computational and experimental forensics characterization of weapons-grade plutonium produced in a thermal neutron environment (Osborn J. 2018b)

Following on to the previous work by Osborn as discussed above, this document focused on categorizing the plutonium produced in the University of Missouri Research Reactor (MURR) to determine if the simulated material was a reasonable representation to the different reactor types of interest. This involved irradiation of material, radiochemical analysis, modeling and simulation of the irradiation, and comparison of calculated and measured data. The methods developed by Osborn showed that reactor physics models could be used to create isotopic data that is representative of physical samples from irradiation and encompassed the fission products and actinides. This indicates that data to develop the classification algorithm does not have to only include physical samples but can also utilize reactor physics model data that is generated and validated against known operating conditions for a specific reactor type.

Operator declaration verification technique for spent fuel at reprocessing facilities (Charlton W., 2000)

One of the main references to the nuclear forensics methodology development paper was this article. This paper came early in the development of nuclear forensics techniques and processes. While the focus was on noble gas collection and analysis, this emphasizes the need for a measurement database, high fidelity models that have been validated, sample measurements with low uncertainty, and data analysis techniques

specific to the problem. Charlton's work was focused on verifying declared information at reprocessing facilities, but the techniques used statistical methods for comparing the sample data to the modeled data. The approach was also identified to be computationally intensive with the expectation that future work could find other methods to reduce the complexity of the database or method.

Assessing confidence in inferring reactor type and fuel burnup: A Markov chain Monte Carlo approach (Burr T., 2005)

The Maximum likelihood method identified in the two previous articles was referenced and evaluated in this article. The purpose was to build upon the maximum likelihood model by using Markov Chain Monte Carlo to simulate observations from the posterior distribution for reactor type and burnup. This adds in the impacts of errors and computer code uncertainties in the modeled data. Errors in measurement were increased by factors of 2, 5, and 10, as a method to evaluate how much error could be tolerated in the MCMC approach. Plutonium Signatures work did similar things to attempt to quantify the effect of measurement uncertainty. Burr's work focused on fission products, notably krypton and xenon, for reactor type identification and utilized 4 different reactor types (PWR, BWR, CANDU, and Liquid Metal Fast Breeder Reactor), with 3 different moderator types (light water, heavy water, and none). The summary identified that incorporating uncertainty in the models and sample data was important, as it affected the results and future efforts should partition the total random and systematic error variances into their respective components using modern experimental design customized for exploring the impact of computer model inputs.

DATA ANALYSIS

Statistical methods involve the use of mathematical techniques, formulas, methods, and data analysis for the purpose of attaining information from raw data. These methods do not have to be complex but can utilize advanced mathematical techniques to find trends and correlations in the dataset. For the work presented in this dissertation, the analytical methods employed started basically with the graphing of the dataset and trying to determine trends in the data, evolved into linear regression of the data, and finished with multivariate regression analysis of the data to capture the correlation in the plutonium isotopic data.

Graphical Data Analysis

One of the easiest methods to analyze data is to plot it to determine if there are any visible relationships. Figure 1 shows calculated data for a variety of reactor types with some at different moderator temperatures. These graphs developed to give a quick estimate of reactor type and to identify potential trends in the data. Unfortunately, these graphs do not capture the uncertainty in the data, which can give a false sense of categorization when comparing data measurements to the plots.

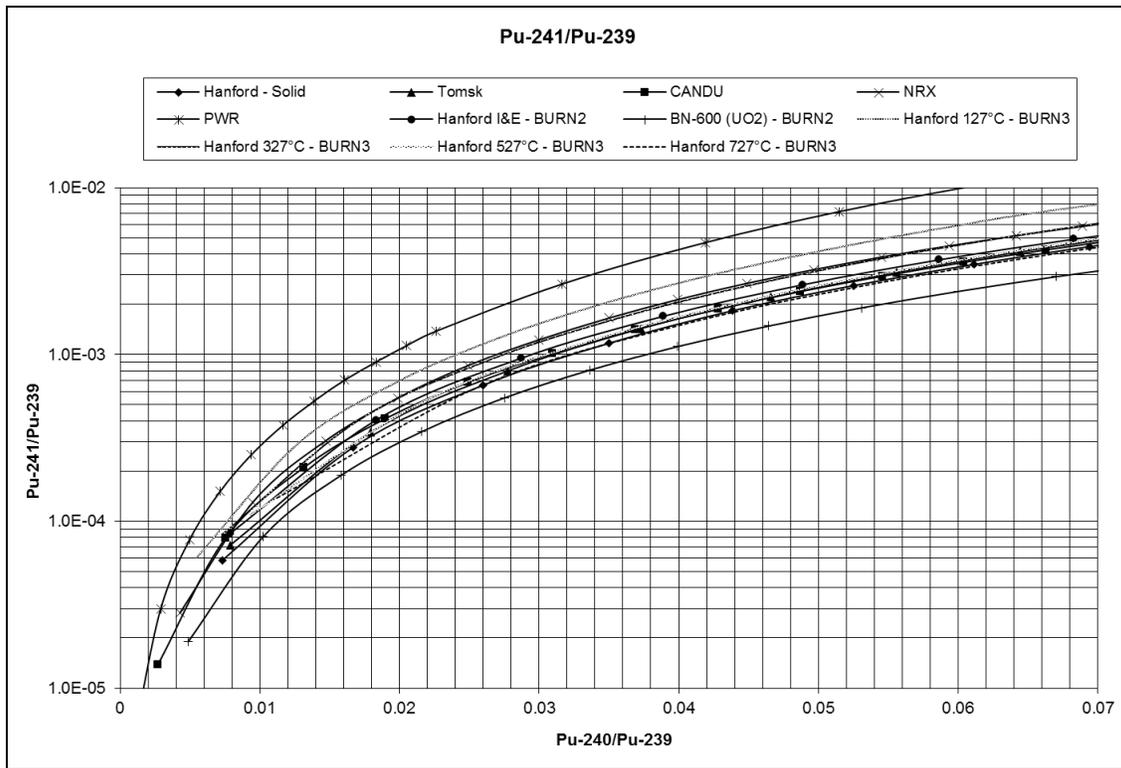


Figure 1: Comparison of $^{241}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ for Different Reactor Types as a Method of Graphical Analysis

Statistical Regression

When evaluating a dataset, it can be useful to determine if there is a relationship between a dependent variable and one or more independent variables. Linear regression can be used when the data has a linear relationship and will find a straight line that has the best fit of bivariate data, in this case how specific plutonium ratios compare to each other. To calculate the least squares regression line for a dataset, the following equation is used:

$$Y = B_0 + B_1X$$

Where:

- B_0 is a constant (defined below)

- B_1 is the regression coefficient (defined below)
- X is the value of the independent variable
- Y is the predicted dependent variable value.

While there are many computational tools that can calculate linear regression of a dataset, the equation to solve for B_0 and B_1 are:

$$B_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$B_0 = \bar{y} - B_1\bar{x}$$

Where:

- x_i is the x value of the i^{th} observation
- \bar{x} is the mean of all x values
- y_i is the y value of the i^{th} observation
- \bar{y} is the mean of all y values

The regression line generated has the properties of passing through the mean X and Y values, B_0 is the y intercept of the regression line, and B_1 is the slope of the line.

Additionally, the coefficient of determination (R^2), also known as the “R-squared value” is an important parameter to calculate for a linear regression. R^2 is the proportion of the variance in the dependent variable that is predictable from the independent variable. The formula to calculate the value for a linear regression model with one independent variable is:

$$R^2 = \left(\frac{\frac{1}{N} \sum [(x_i - \bar{x})(y_i - \bar{y})]}{(\sigma_x * \sigma_y)} \right)^2$$

Where:

- N is the number of observations

- x_i is the x value of the i^{th} observation
- \bar{x} is the mean of all x values
- y_i is the y value of the i^{th} observation
- \bar{y} is the mean of all y values
- σ_x is the standard deviation of x
- σ_y is the standard deviation of y

Linear regression in this analysis was used to predict an isotope ratio (either $^{238}\text{Pu}/^{239}\text{Pu}$, $^{241}\text{Pu}/^{239}\text{Pu}$, or $^{241}\text{Pu}/^{239}\text{Pu}$) based on the given value for $^{240}\text{Pu}/^{239}\text{Pu}$. However, linear regression was not found to be useful for the purposes of reactor classification. To use the linear regression analysis, each reactor type would have a different regression and new values of isotope ratios would be compared against the regressions to find a likely match. As the data for different reactor types overlap to a significant degree (see Figure 1), the ability to identify a reactor type with certainty against a single parameter is difficult.

Multiple Regression Analysis

Sometimes the prediction made using a single independent variable is not sufficient to make an accurate determination, or there is additional data available that can be used to make a better prediction. With the complexities of reactor operation and analysis, it was expected that linear regression would not have the necessary accuracy for reactor attribution. In this case, there is more than one independent variable to be considered, therefore multiple regression analysis can extend from the linear regression to increase the ability to classify reactor types from the available data. The theory section of this dissertation will expand on the complexity of reactor operation, reaction rates, and plutonium production.

For multiple regression, we expand on the linear relationship of a single independent variable by adding more. In this example case, we have two:

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_nX_n$$

Where:

- B_0 to B_n are the regression coefficients that need to be estimated
- X_1 to X_n are the independent variables
- Y is the predicted dependent variable value.

When dealing with only one independent variable, as identified in the previous subsection, the problem can be described graphically by plotting the independent variable to the dependent variable. As the number of independent variables increases, the number of dimensions to graphically represent the data also increases. In the case of two independent variables, the regression equation represents a surface that is described by the mean values of Y at various combinations of X_1 and X_2 . The simplest representation of this is a plane in three-dimensional space. The following equation represents the least-squares solution that gives the best fit:

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - B_0 - B_1X_{1,i} - B_2X_{2,i} - \dots - B_nX_{n,i})^2$$

Given the above equation, the next step is to find the regression coefficients that satisfy the least-squares criterion. For the actual analysis, this was done in R, however, for the purposes of understanding the theory, this is explained with the use of matrix algebra.

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

$$B = \begin{bmatrix} B_0 \\ B_1 \\ B_2 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & X_{1,1} & X_{1,2} \\ 1 & X_{2,1} & X_{2,2} \end{bmatrix}$$

Given the above matrices, the multiple regression equation is:

$$Y = XB$$

And the least-squares normal equations are expressed as:

$$X'Y = X'XB$$

And solving for the regression coefficients:

$$(X'X)^{-1}X'Y = (X'X)^{-1}X'XB = B$$

Where:

- X' is the transpose of X
- $(X'X)^{-1}(X'X)=I$, or the identity matrix

THEORY

This dissertation is based upon measurements and calculations of the production and transmutation of nuclides in nuclear reactor fuel during irradiation. Nuclear reactors are an excellent source of large quantities of neutrons, which are necessary to produce plutonium in quantities greater than a laboratory scale. At a high level, the nuclear fuel is exposed to neutrons which can create fission interactions that liberate more neutrons, or the neutrons can be absorbed in other materials. Plutonium is produced when a neutron is absorbed in the non-fissile isotope ^{238}U and undergoes a couple of radioactive decays. Multiple factors of the reactor type and operational parameters impact the probability that certain reactions will occur and ultimately affect the isotopics of the plutonium that is produced. Nuclear forensics and ultimately the attribution of plutonium uses these differences in plutonium isotopics to discriminate against the different reactor types of material origin.

While it would be ideal to have a single comprehensive measurement database of all materials and isotopes of interest, the sheer volume of measurements that would need to be completed is unrealistic, and most of these materials that would need to be measured are unavailable for a variety of reasons. To facilitate the addition of data into the classification model, reactor physics calculations must be performed to determine the effect of varying operating parameters on the output isotopics. Of course, as additional measured data is generated or discovered, that data can be ingested into the classification model for increased performance.

This work focused on the production of different isotope ratios for the purpose of constructing a classification algorithm. In this model, the primary focus is on actinides with an emphasis on plutonium isotopics consistent with *Luksic, et al* which showed that

isotopic ratios of plutonium were correlated to the reactor type of origin. (Luksic A., 2012) Luksic's work also showed that there could be areas of overlap between reactor types depending on the range of reactor operating conditions considered. The specific operating parameters that were identified in the NEA report *Spent Nuclear Fuel Assay Data for Isotopic Validation* (NEA, 2011) report including: initial uranium fuel isotopics, moderator type and temperature, reactor power level, irradiation time, and void coefficient for BWRs, will be discussed for their impact on plutonium isotopics.

Nuclear Interactions and Isotope Production

The calculation of individual nuclide concentrations is dependent upon their formation and removal due to radioactive disintegration and transmutation via nuclear interactions. For the purposes of plutonium production, uranium fuel used in nuclear reactors is transmuted to plutonium during irradiation and then chemically separated from the reactor fuel. The neutron absorption cross sections play a major role in determining the probability of interaction to create the desired isotope. In many instances there are competing probabilities which can include neutron absorption or neutron induced fission.

There are many different types of nuclear interactions that result in both the operation of nuclear reactors and the production of plutonium. The following subsections describe the general types of neutron interactions of interest in nuclear reactors and isotope production. With a general understanding of these interactions, additional information on neutron cross sections and will be introduced to further demonstrate the importance of specific reactor operating parameters and conditions on plutonium production and ultimately classification of a reactor type of origin for a source of plutonium.

NUCLEAR INTERACTIONS

Scattering Interactions

There are two types of interactions where a neutron with an initial kinetic energy interacts with a nucleus and leaves with a lower kinetic energy. The target nucleus and the neutron do not change, but the amount of kinetic energy each has after the interaction is different. These two types of interaction: elastic scattering and inelastic scattering are discussed in more detail.

Elastic Scattering Interactions

An elastic scattering collision looks like a billiard ball collision. In this case, a neutron collides with a nucleus, transfers some energy to it, and bounces off in a different direction. There can be instances where the nucleus absorbs the neutron and then reemits it, while conserving kinetic energy. The amount of initial energy lost depends on the angle of interaction. The target nucleus gains the amount of energy lost by the neutron as an increase in kinetic energy and. Figure 2 shows an illustration of an elastic scattering interaction. A nucleus that is on the same relative size of a neutron will be more effective at energy transfer than a larger nucleus.

Inelastic Scattering Interactions

Inelastic scattering is where a neutron interacts with a nucleus and is temporarily absorbed, forming a compound nucleus. This creates an excited state and may de-excite by emitting another neutron of lower energy kinetic energy with a gamma, which takes the remaining energy. This interaction generally happens only with high-energy neutrons interacting with heavy nuclei. For inelastic scattering, the kinetic energy of the system is

not conserved, as energy is reduced from the emitted gamma. Figure 3 shows an illustration of an inelastic scattering interaction.

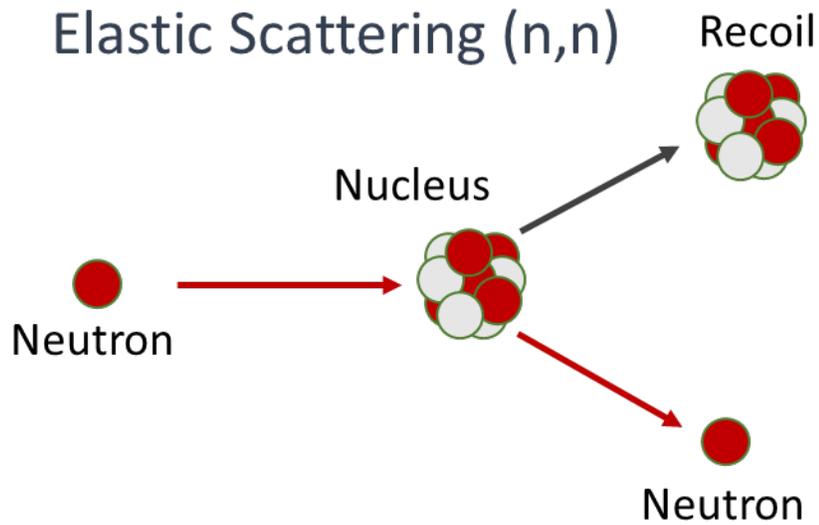


Figure 2: Elastic scattering interaction with a neutron and a nucleus

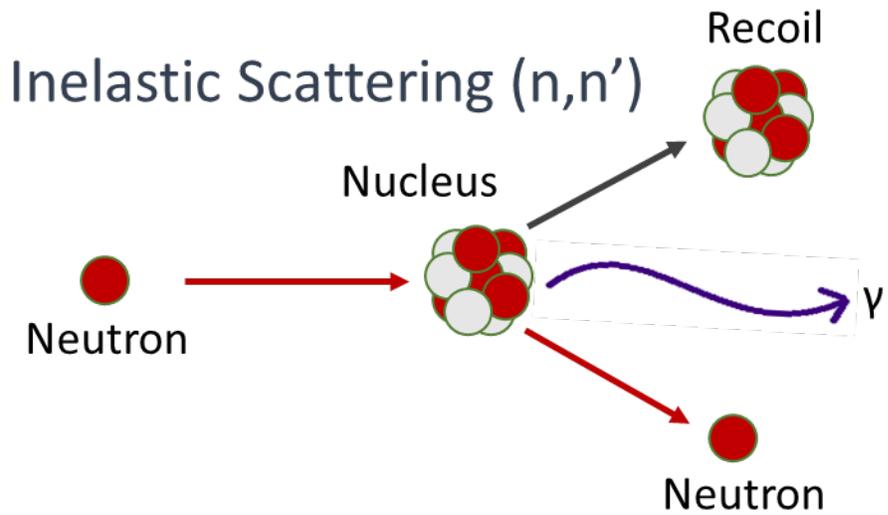


Figure 3: Inelastic scattering interaction with a neutron and a nucleus, with a gamma emitted

Transmutation

In a transmutation interaction, a target nucleus will absorb a neutron and form a compound nucleus. The compound nucleus will then emit a charged particle, which can be either a proton or alpha particle. Depending on the energy of the incoming neutron, a gamma may also be emitted when the final nucleus deexcites. Since the interaction results in a different element than the initial nucleus, we refer to this interaction as transmutation. The (n,p), neutron-proton transmutation, and the (n, α) neutron-alpha transmutation will be discussed in further detail.

(n,p) Neutron-Proton Transmutation

This reaction occurs when a neutron enters a nucleus and then a proton exits the nucleus. Figure 4 shows an example where a neutron enters an ^{16}O nucleus, and then the compound nucleus ejects a proton to become ^{16}N . In this type of interaction, the total number of nucleons in the nucleus stays the same, 16 for this example, but the number of protons is reduced by one, resulting in a different element. One way to notate this type of reaction is by writing the target nucleus (^{16}O), then the reaction type (n,p), and then the resulting nucleus (^{16}N) to give the final form of $^{16}\text{O}(n,p)^{16}\text{N}$. Most (n,p) reactions have threshold neutron energies that are required for the reaction to take place. If the incoming neutron is below this energy, the reaction will not happen.

Transmutation (n,p)

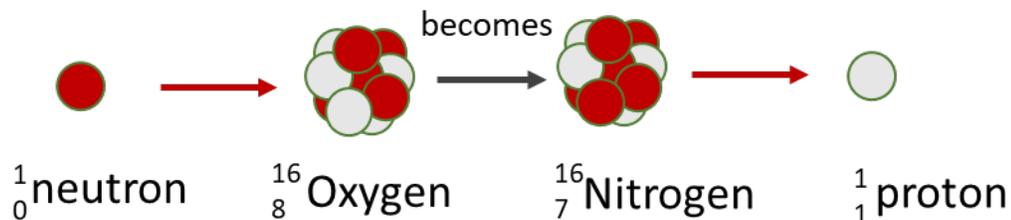


Figure 4: Neutron-proton transmutation interaction: ${}^{16}\text{O}(n,p){}^{16}\text{N}$

(n, α) Neutron-Alpha Transmutation

This reaction occurs when a neutron enters a nucleus and then an alpha particle exits the nucleus. Figure 5 shows an example where a neutron enters a ${}^{10}\text{B}$ nucleus, and then the compound nucleus ejects an alpha particle to become ${}^7\text{Li}$. In this type of interaction, the total number of nucleons in the resulting nucleus is reduced by three, with the number of protons is reduced by two, resulting in a different element. Following the same notation as the (n,p) reaction, we start with the target nucleus (${}^{10}\text{B}$), then the reaction type (n, α), and then the resulting nucleus (${}^7\text{Li}$) to give the final form of ${}^{10}\text{B}(n,\alpha){}^7\text{Li}$.

Transmutation (n, α)

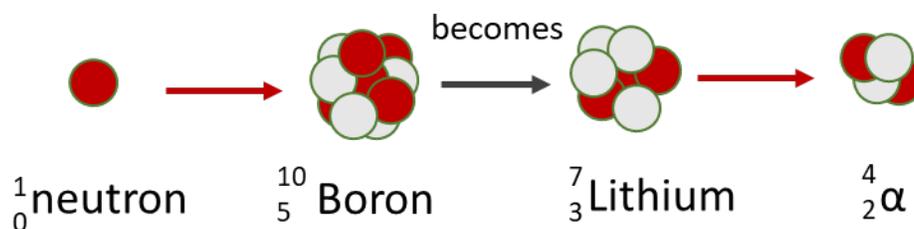


Figure 5: Neutron-alpha transmutation interaction: ${}^{10}\text{B}(n,\alpha){}^7\text{Li}$

Capture Interactions

One of the more common nuclear reactions is radiative capture. This interaction is where a nucleus absorbs a neutron and the compound nucleus that is created emits only a gamma. The resulting nucleus is the same element, but the number of neutrons increases by one, changing the isotope. Figure 6 shows an example where a neutron enters a ^{239}Pu nucleus, and then the compound nucleus ejects a gamma as it becomes ^{240}Pu . This type of reaction is more likely to occur with low energy, or thermal, neutrons.

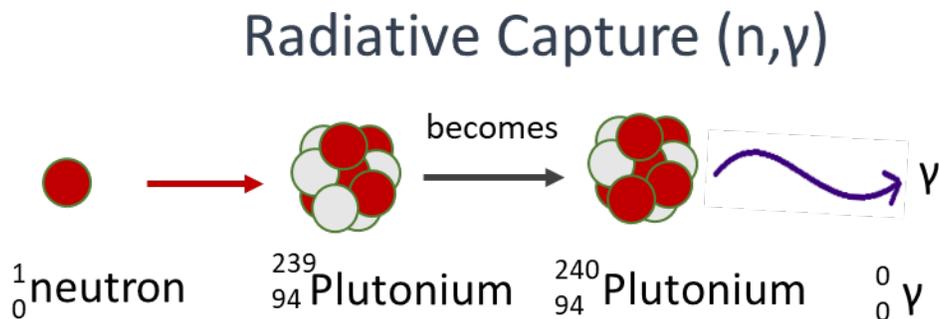


Figure 6: Neutron-gamma, radiative capture interaction: ${}^{239}\text{Pu}(n,\gamma){}^{240}\text{Pu}$

Fission

For the purposes of nuclear reactor operation and plutonium production, the fission reaction is the most important. The fission reaction is where a neutron interacts with a nucleus and the compound nucleus splits into two or more smaller nuclei and releases neutrons, gammas, and energy. Fission is considered a transmutation interaction because the resulting product nuclei are a different element than the original target. One of the important aspects of this reaction is that usually two or more neutrons are emitted during this process. This allows for the fission interactions to be chained together, or self-sustaining, resulting in the operation of a nuclear reactor. These fission neutrons are also important in the production of plutonium as any that are captured in uranium can contribute

to production. Figure 7 shows an example of a fission reaction, where a neutron interacts with a nucleus, induces a fission event, and results in two fission fragments being produced along with three neutrons and a couple of gammas.

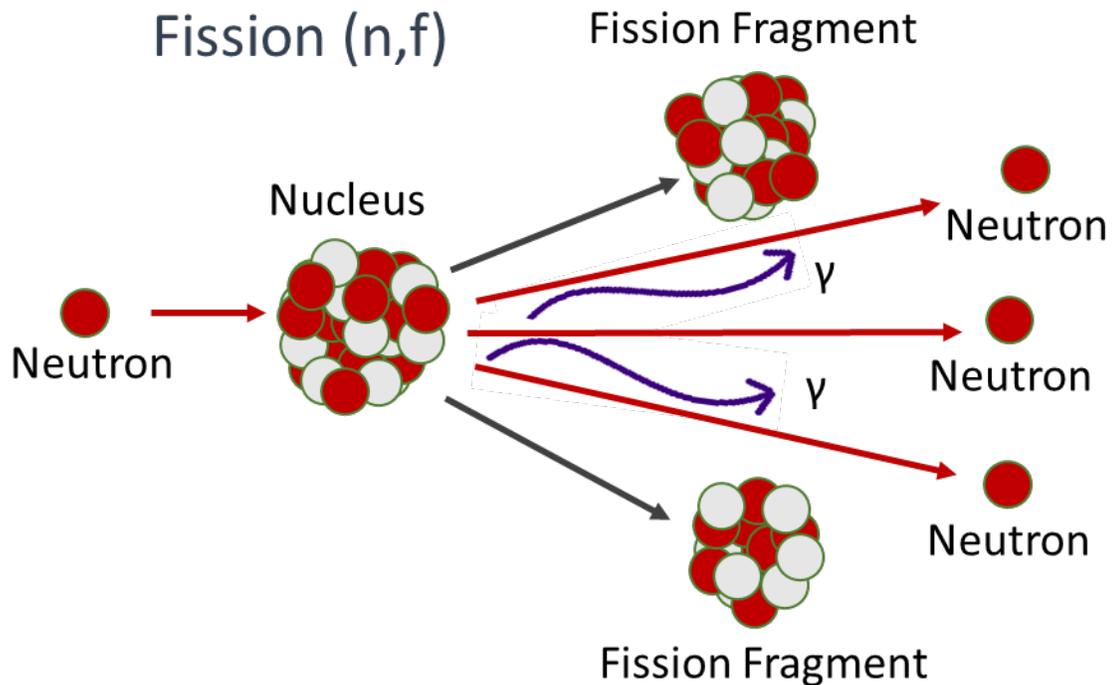


Figure 7: Example of a neutron induced fission interaction

NUCLEAR CROSS SECTIONS

In a nuclear reactor, the neutron reactions previously discussed will all have different probabilities of occurrence, and these probabilities largely depend on the energy of the incident neutron. The concept of neutron cross sections is used to describe the likelihood of the specific interactions occurring. The two main types of cross sections that are of concern in a nuclear reactor are the scattering and absorption cross sections. These reactions line up with the interactions described previously, with the scattering cross sections

applying to the elastic and inelastic scattering, and the absorption reactors relating to the transmutation, capture, and fission interactions.

The standard unit for measuring the cross section is the barn, which is in units of area and is equal to $1 \times 10^{-24} \text{ cm}^2$ or $1 \times 10^{-28} \text{ m}^2$. In general, the larger the cross section, the greater the probability of a neutron interaction with the nucleus. Figure 8 shows the four major interaction cross sections for ^{235}U , including total, scattering, capture, and fission, with the data for this figure obtained from the National Nuclear Data Center ENDF/B-VII.1 libraries. Each of these interactions has a probability of occurring at any energy, however, depending on the energy of the incident neutron, and specific interaction will have a higher probability of occurrence. For example, a scattering interaction is most probable at energies greater than 10 keV, while a fission interaction has the most probable occurrence at energies less than $1 \times 10^{-6} \text{ MeV}$.

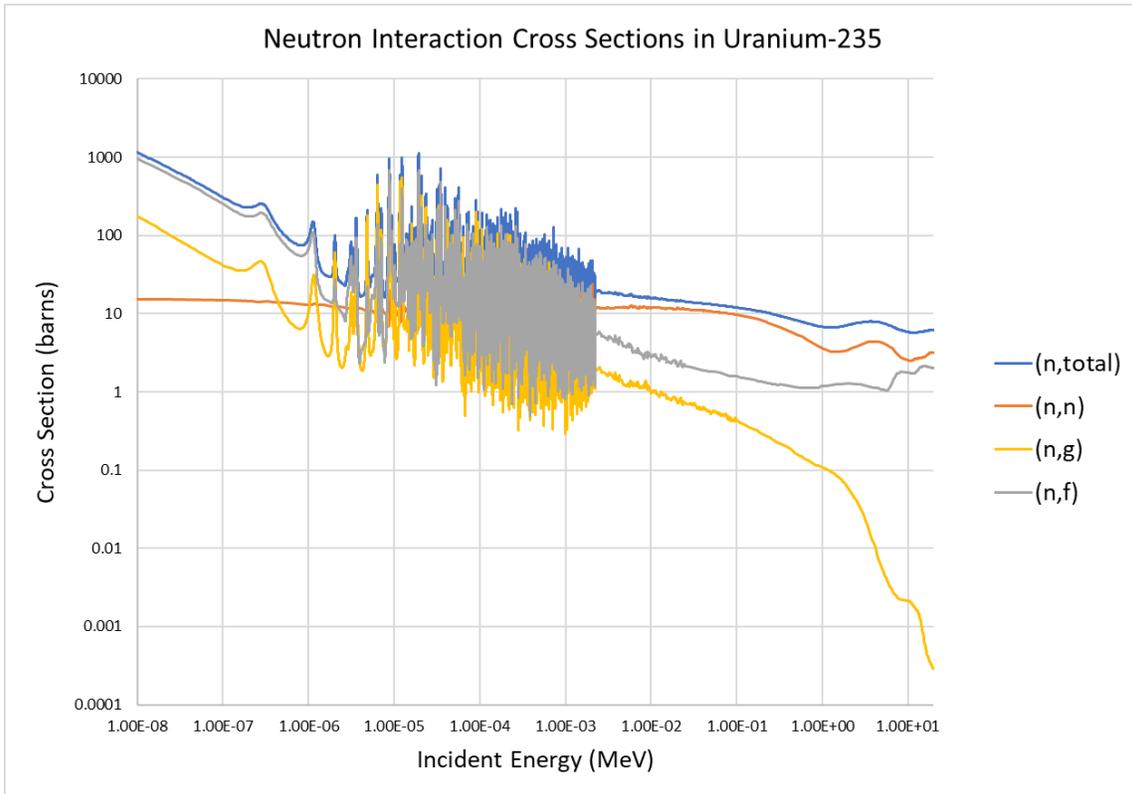


Figure 8: Total, Scattering, Capture, and Fission Neutron Cross Sections in ^{235}U (Chadwick M., 2011)

Energy Regions

Since cross sections are a function of energy, accounting for the energy is important for accurately modeling the behavior of nuclear reactors. The energy spectrum of the reactor will have a noticeable impact on plutonium isotopics. Figure 8 shows the continuous cross section data for ^{235}U , but to reduce the complexity of calculations, the cross sections can be consolidated into discrete energy regions with similar behavior. Depending on the fidelity of the nuclear cross section data, the confidence of the neutron energy, and the computational resources available, the energy range can be divided into multiple regions. As an example, energy can be broken down into the following 10 regions

identified from *nuclear-power.com*:

- **Cold Neutrons** (0 eV to 0.025 eV). Neutrons in thermal equilibrium with very cold surroundings such as liquid deuterium. This spectrum is used for neutron scattering experiments.
- **Thermal Neutrons** (0.025 eV). Neutrons in thermal equilibrium with a surrounding medium. Most probable energy at 20°C (68°F) for Maxwellian distribution is 0.025 eV (~2 km/s). This part of the neutron's energy spectrum constitutes the most important spectrum in thermal reactors.
- **Epithermal Neutrons** (0.025 eV to 0.4 eV). Neutrons of kinetic energy are greater than thermal. Some reactor designs operate with an epithermal neutron spectrum. This design allows reaching a higher fuel breeding ratio than in thermal reactors.
- **Cadmium Neutrons** (0.4 eV to 0.5 eV). Neutrons of kinetic energy below the cadmium cut-off energy. One cadmium isotope, ^{113}Cd , absorbs neutrons strongly only if they are below ~0.5 eV.
- **Epicadmium Neutrons** (0.5 eV to 1 eV). Neutrons of kinetic energy above the cadmium cut-off energy. These neutrons are not absorbed by cadmium.
- **Slow Neutrons** (1 eV to 10 eV).
- **Resonance Neutrons** (10 eV to 300 eV). The resonance neutrons are called resonance for their special behavior. At resonance energies, the cross-sections can reach peaks more than 100x higher than the base value of the cross-section. At these energies, the neutron capture significantly exceeds the probability of fission. Therefore, it is very important (for thermal reactors) to quickly overcome this range of energy and operate the reactor with thermal neutrons increasing the probability of fission.
- **Intermediate Neutrons** (300 eV to 1 MeV).
- **Fast Neutrons** (1 MeV to 20 MeV). Neutrons of kinetic energy greater than 1 MeV (~15 000 km/s) are usually named fission neutrons. These neutrons are produced

by nuclear processes such as nuclear fission or (α,n) reactions. The fission neutrons have mean energy (for ^{235}U fission) of 2 MeV. Inside a nuclear reactor, the fast neutrons are slowed down to thermal energies via a process called neutron moderation.

- **Relativistic Neutrons** (greater than 20 MeV). (Nuclear Power, 2022)

For the purposes of reactor physics computer codes and calculations, this fine energy structure is not necessary. Looking at Figure 8, notice the cross section can be visually separated into three different regions. Figure 9 shows the neutron interactions cross sections for ^{235}U previously seen in Figure 8 with three different cross section regions identified: the thermal region, the resonance region, and the fast region. For the purposes of this research the three regions are sufficient to describe the necessary interactions that drive the production of plutonium and its transmutation into the higher plutonium isotopes.

The regions are divided into the following energy groups:

- **Thermal Neutrons** (0 eV to 1 eV).
- **Resonance Neutrons** (1 eV to 0.5 MeV).
- **Fast Neutrons** (greater than 0.5 MeV).

The following subsections will describe the energy range and the overall significance of each region in relation to reactor physics calculations.

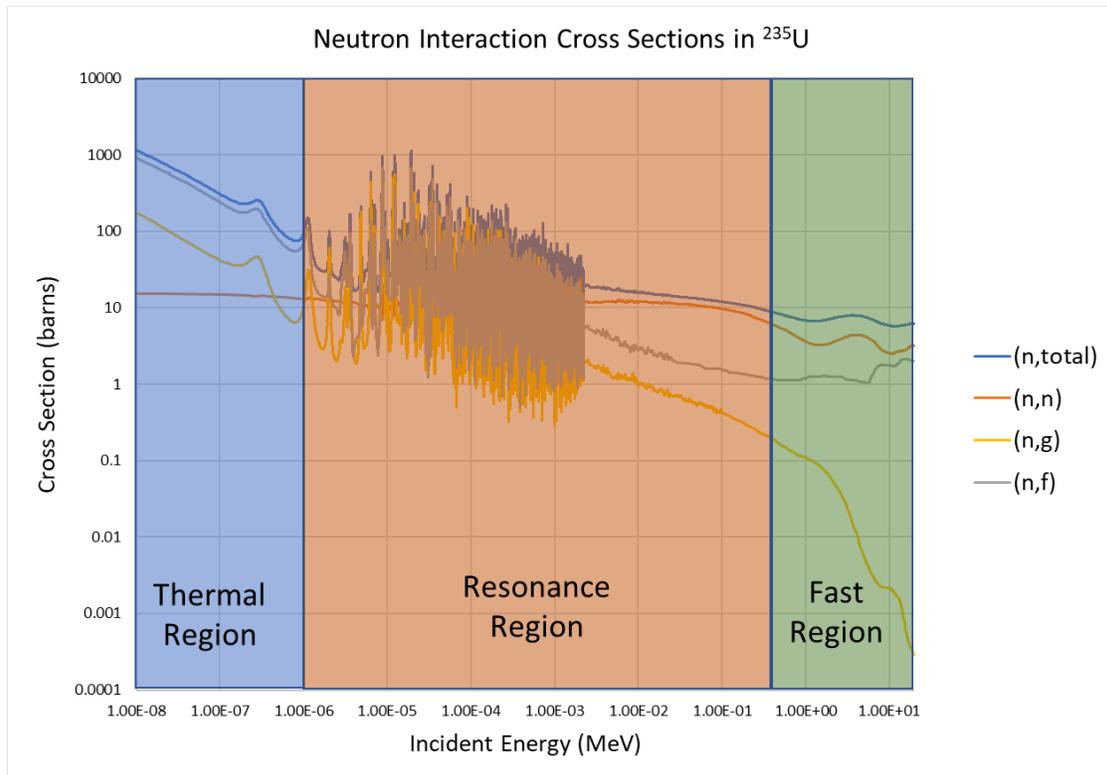


Figure 9: Neutron Cross Sections in ^{235}U with the specific regions identified (Chadwick M., 2011)

Thermal Region

For this research, the thermal region is defined as a neutron energy of less than 0.1 eV. This includes neutrons that are in thermal equilibrium with its surrounding medium, having a temperature of approximately 20 C. This energy region is important, as shown in Figure 9, because the fission cross section is the dominant interaction in this region. When operating a reactor design with natural or low enriched uranium, neutrons must be “thermalized” to maintain a fission rate that supports the chain reaction. Since neutrons are “born” from fission with high energies (in the MeV range), multiple scattering interactions must occur for the fission neutrons to transfer the energy to reach the thermal region.

Resonance Region

For this research, the resonance region is defined as neutron energy between 1 eV and 0.5 MeV. As noted in the description of the ten energy groups, the resonance energies represent the region where the neutron capture cross section are greater than the fission cross section. This energy region is important for nuclear reactor operation, plutonium production, and plutonium isotopics because of the competing need for neutrons to transition through the range to the thermal region without being absorbed to sustain the chain reaction, while neutrons absorbed in ^{238}U are necessary for plutonium production.

Fast Region

For this research, the fast region is defined as neutron energy greater than 0.5 MeV. These neutrons are usually produced by fission but can also be produced by other nuclear interactions (n,2n; α ,n; etc.). For thermal reactors, neutrons that have energies in the fast region need to be moderated and transfer energy to reach the thermal region to sustain the fission chain reaction.

Figure 8 only represents three interaction types and the total cross section for only the ^{235}U isotope. Nuclear reactors and nuclear fuel are comprised of many different fissionable isotopes, all of which will have competing cross sections for neutrons when the reactor is operating. The precursor to determining the origin of irradiated nuclear material is understanding the different physical processes that occur during irradiation and then leverage those in mathematical and statistical models. These models will be based on measurement data obtained from irradiated fuel rods and validated reactor physics models to calculate data to fill in gaps.

CROSS SECTIONS AND REACTOR OPERATION

Most of the nuclear reactor fuel in the world, and all the fresh fuel used for plutonium production is primarily uranium. Table 1 shows the natural abundance of uranium isotopes and the cross sections

Isotope	Natural Abundance	Thermal Absorption cross-section (b)	Resonance Absorption cross-section (b)	Thermal Fission cross-section (b)
^{234}U	0.0054%	106	700	0.067
^{235}U	0.7204%	185.72	143.76	585
^{236}U	0.0%	5.1	360	0.04
^{238}U	99.2742%	2.68	277	0.000005

Table 1: The natural abundance of U isotopes and their thermal capture, resonance radiative capture, and fission cross-sections (Chart of the Nuclides, 17th Edition).

Depending on the interaction, the amount and corresponding isotopic distribution of the plutonium produced will vary. Figure 10 shows the uranium fission and absorption neutron cross sections and Figure 11 shows the same cross sections for plutonium. The data for these figures was obtained from the National Nuclear Data Center ENDF/B-VII.1 libraries. (Chadwick M., 2011) Note that there is a relatively high thermal fission cross section for ^{235}U , ^{239}Pu , and ^{241}Pu and similar resonance absorption cross sections for all the isotopes. It is important to define some terms relating to the fission of isotopes: Fissionable, Fissile, and Fertile.

- Fissionable nuclides will undergo neutron induced fission but may or may not be able to sustain a chain reaction and includes most uranium and plutonium isotopes. As an example, ^{238}U is fissionable, but it is unable to sustain a chain reaction.
- Fissile nuclides will undergo neutron induced fission with thermal neutrons. These isotopes can sustain a nuclear chain reaction and include ^{233}U , ^{235}U , and ^{239}Pu .

- Fertile nuclides can be transmuted into fissile nuclides by neutron capture. These include ^{238}U and ^{232}Th , which are transmuted into ^{239}Pu and ^{233}U , respectively.

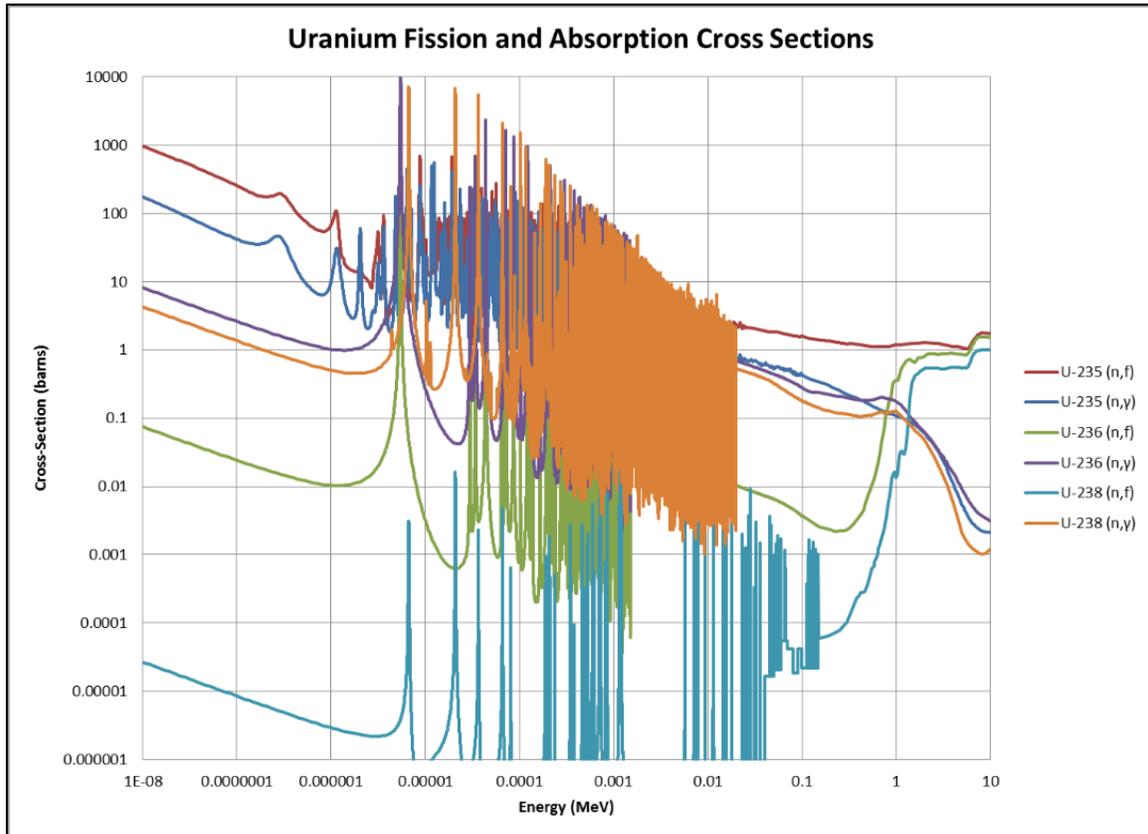


Figure 10: Fission and Absorption Neutron Cross Sections for Uranium (Chadwick M., 2011)

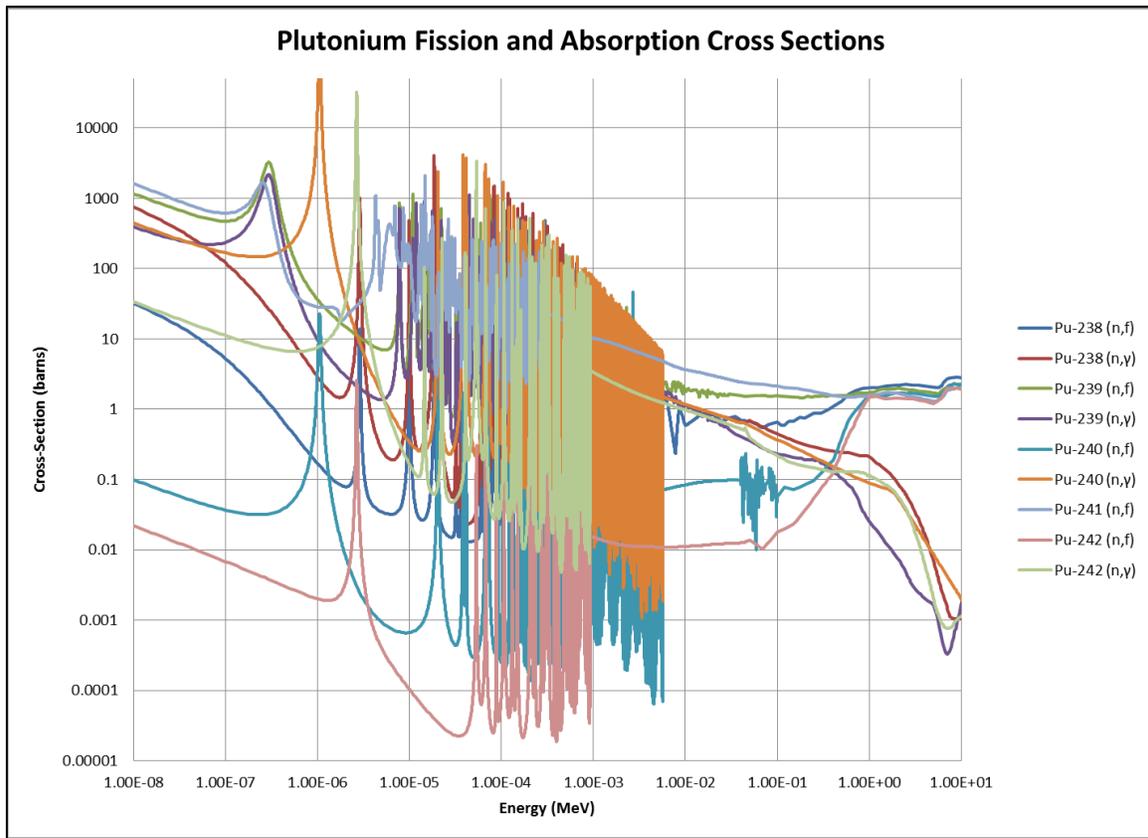


Figure 11: Fission and Absorption Neutron Cross Sections for Plutonium (Chadwick M., 2011)

The previous figures show that for a fissile nuclide, the likelihood of fission is higher for lower neutron energies. As an example, a 1 eV neutron is about 300 times more likely to induce a ^{235}U fission than a 1 MeV neutron. Figure 12 shows the prompt neutron energy spectrum. Prompt neutrons are neutrons that are released at the instant of fission, these are different than delayed neutrons, which are released in varying times after the fission has occurred. Note that the most probable energy of a prompt fission neutron is 0.73 MeV. Prompt fission neutrons are born in the fast energy region and must be moderated to transfer energy to other materials to get down into the thermal energy region.

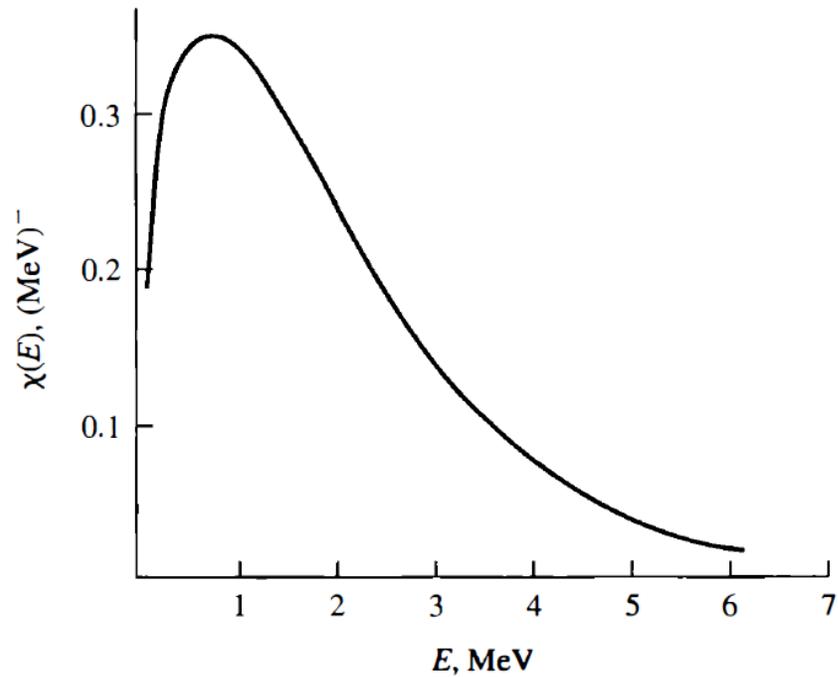


Figure 12: Prompt Neutron Energy Spectrum (Lamarsh J., 2001)

Reactor Operation

To maintain a nuclear chain reaction to operate a nuclear reactor, a neutron balance must occur that keeps the chain reaction going. From the neutron's perspective, there are a few things that can happen once it is "born" from a fission event:

1. They can scatter off other nuclei and transfer energy
2. They can be captured in non-fission interactions
3. They can leak out of physical boundary of the reactor
4. They can be captured and cause a fission event, releasing more neutrons

To successfully operate a nuclear reactor, these events and interactions must be balanced so that the chain reaction continues. Figure 13 shows a representative neutron life cycle in a nuclear reactor with the types of interactions described in greater detail.

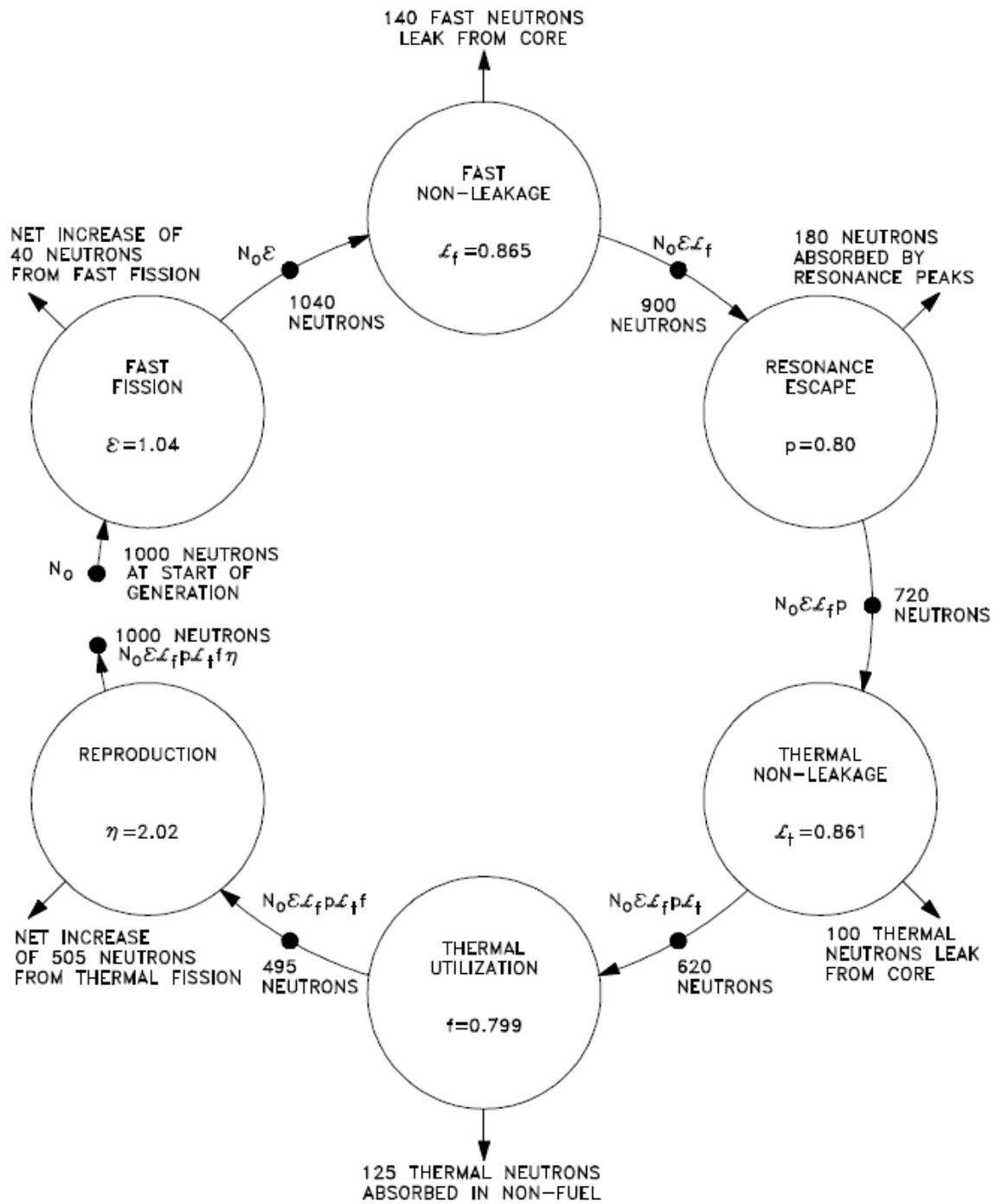


Figure 13: Neutron Life Cycle with $k_{eff} = 1$ (DOE, January 1993)

The figure gives a notional generation of 1000 neutrons, where a generation represents the number of neutrons that the system starts with and then after all the neutrons interact in events previously described, how many neutrons start the next generation. For a nuclear reactor to maintain its chain reaction, the number of neutrons in each generation must be the same. From Figure 13, there are six things that can happen to the neutron in the reactor: a fission interaction at fast energy, leaking out of the core at fast energy, absorbed in the resonance energy region, leaking out of the core at thermal energies, absorption at thermal energies without resulting in fission, and fission at thermal energies. The reactor designer must take these reactions into account when designing the reactor. Additionally, different fuel and moderator types will have an impact the probabilities of these events occurring. These differences in cross sections and interaction probabilities can be observed in the different isotope production and isotopic ratios and can ultimately be used to determine a what type of reactor was used.

The next section will identify the production of the different plutonium isotopes. Note that this is a simplification of the complex reactor physics going on in the reactor to produce plutonium isotopes. This section is not intended to solve the differential equations for plutonium isotope production but give a general description of the processes involved.

ISOTOPE PRODUCTION

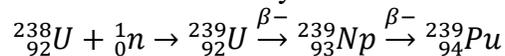
This section focuses on the production of plutonium isotopes through the various interactions described in the previous sections. As noted above, nuclear reactors are the only source of a sufficient number of neutrons to produce quantities of plutonium beyond a laboratory scale. Any reactor that uses uranium fuel will produce plutonium, but some reactors are better at producing it than others. Ultimately, the type of reactor and the operating conditions will affect the plutonium isotopics that are produced. The goal is to

have a reactor design that achieves the purpose of the reactor, which can include power production, nuclear research, or isotope production. The plutonium that is produced during the reactor operation will have different characteristics that can be leveraged in the attribution of a specific reactor type. Utilizing the Bateman Equations, production, loss, and decay are calculated to determine isotopic production.

Bateman Equations

²³⁹Pu Production

The following equation identifies the production pathway for ²³⁹Pu which includes a neutron absorption in ²³⁸U and two beta decays to reach ²³⁹Pu.



The first beta decay to ²³⁹Np has a half-life of 23.5 minutes and the second beta decay to ²³⁹Pu has a half-life of 2.36 days. Since the process of irradiation and reprocessing of the fuel for the purpose of plutonium production is on the order of weeks to months, we can assume that the radiative capture of a neutron by ²³⁸U is converted in to ²³⁹Pu and the time dependent amount of plutonium is basically determined by:

$$\begin{aligned} \frac{dN^{Pu239}}{dt} = & \phi\sigma_c^{U238}N^{U238}(t) - \lambda_{Pu239}N^{Pu239}(t) - \phi\sigma_c^{Pu239}N^{Pu239}(t) \\ & - \phi\sigma_f^{Pu239}N^{Pu239}(t) \end{aligned}$$

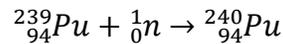
Where:

- N(t) is the function the number density with respect to time
- t is the irradiation time
- ϕ is the flux within the reactor
- σ_c is the radiative capture cross-section
- σ_f is the fission cross-section
- λ is the decay constant

For the timescale we are referring to for irradiation and reprocessing, the decay term is negligible for ^{239}Pu , though it was included for completeness, and the production is dependent on the radiative capture rate of ^{238}U and the losses are determined by the radiative capture and fission capture rate of the produced ^{239}Pu .

^{240}Pu Production

The following equation identifies the production pathway for ^{240}Pu which includes a neutron absorption in ^{239}Pu .



The production of ^{240}Pu is dependent on the radiative capture of a neutron by ^{239}Pu and is basically determined by:

$$\frac{dN^{Pu240}}{dt} = \phi\sigma_c^{Pu239}N^{Pu239}(t) - \phi\sigma_c^{Pu240}N^{Pu240}(t)$$

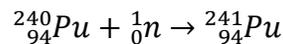
where

- $N(t)$ is the function the number density with respect to time
- t is the irradiation time
- ϕ is the flux within the reactor
- σ_c is the radiative capture cross-section

As in the ^{239}Pu case, the decay term is negligible, and we are assuming a negligible loss rate from fission so those loss terms are removed from the equation.

^{241}Pu Production

The following equation identifies the production pathway for ^{241}Pu which includes a neutron absorption in ^{240}Pu .



The production of ^{241}Pu is dependent on the radiative capture of a neutron by ^{240}Pu and is basically determined by:

$$\frac{dN^{Pu241}}{dt} = \phi\sigma_c^{Pu240}N^{Pu240}(t) - \lambda_{Pu241}N^{Pu241}(t) - \phi\sigma_c^{Pu241}N^{Pu241}(t) - \phi\sigma_f^{Pu241}N^{Pu241}(t)$$

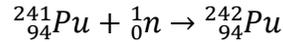
where

- $N(t)$ is the function the number density with respect to time
- t is the irradiation time
- ϕ is the flux within the reactor
- σ_c is the radiative capture cross-section
- σ_f is the fission cross-section
- λ is the decay constant

The production of ^{241}Pu is dependent on the radiative capture rate of ^{240}Pu and the losses are determined by the decay constant, radiative capture, and fission capture rate of the produced ^{241}Pu . With a 14.4-year half-life, beta decay to ^{241}Am is not negligible and must be considered in the production and loss equation. Additionally, the amount of ^{241}Am that is in a plutonium sample can be used to determine specific information about the sample, including the time since the plutonium was separated, or discharged from the reactor. Depending on the age of a plutonium sample, the isotopes for ^{241}Pu and ^{241}Am may be combined into a summed value to approximately indicate the total amount of ^{241}Pu at fuel discharge. Depending on the use of the isotopic information, it may be important to decay correct the plutonium isotopes back to the discharge parameters or having the information on time since discharge or time since last reprocessing can be used to compare against available capabilities to identify or eliminate locations of production.

²⁴²Pu Production

²⁴²Pu production follows the similar pathway of ²⁴⁰Pu. The following equation identifies the production pathway for ²⁴¹Pu which includes a neutron absorption in ²⁴¹Pu.



The production of ²⁴²Pu is dependent on the radiative capture of a neutron by ²⁴¹Pu and is basically determined by:

$$\frac{dN^{\text{Pu}242}}{dt} = \phi\sigma_c^{\text{Pu}241}N^{\text{Pu}241}(t) - \phi\sigma_c^{\text{Pu}242}N^{\text{Pu}242}(t)$$

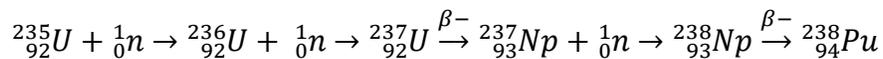
where

- N(t) is the function the number density with respect to time
- t is the irradiation time
- ϕ is the flux within the reactor
- σ_c is the radiative capture cross-section

As in the ²³⁹Pu and ²⁴⁰Pu cases, the decay term is negligible, and we are assuming a negligible loss rate from fission, so those loss terms are removed from the equation.

²³⁸Pu Production

The following equation identifies the production pathway for ²³⁸Pu which includes a neutron absorption in ²³⁵U that does not result in fission, a neutron absorption in ²³⁶U and a beta decay to produce ²³⁷Np followed by another neutron absorption and beta decay to reach ²³⁸Pu.



This pathway produces very small amounts of ²³⁸Pu in comparison to other plutonium isotopes. With the relatively low radiative capture cross section of ²³⁵U, the amount of ²³⁶U

that is produced is very small. Then while working up the chain, the first beta decay to ^{237}Np has a half-life of 6.75 days followed by a radiative capture to ^{238}Np and then a beta decay to ^{238}Pu has a half-life of 2.12 days. Complicating matters further, ^{237}Np , ^{238}Np and ^{238}Pu all have fission cross sections which increase the removal as you move up the chain.

The time dependent amount of plutonium is basically determined by:

$$\begin{aligned} \frac{dN^{Pu238}}{dt} = & \phi\sigma_c^{Np238} \left(\phi\sigma_c^{Np237} \left(\phi\sigma_c^{U236} \left(\phi\sigma_c^{U235} N^{U235}(t) \right) \right) \right) \\ & - \phi\sigma_f^{Np237} \left(\phi\sigma_c^{U236} \left(\phi\sigma_c^{U235} N^{U235}(t) \right) \right) \\ & - \phi\sigma_f^{Np238} \left(\phi\sigma_c^{Np237} \left(\phi\sigma_c^{U236} \left(\phi\sigma_c^{U235} N^{U235}(t) \right) \right) \right) - \lambda_{Pu238} N^{Pu238}(t) \\ & - \phi\sigma_c^{Pu238} N^{Pu238}(t) - \phi\sigma_f^{Pu238} N^{Pu238}(t) \end{aligned}$$

where

- $N(t)$ is the function the number density with respect to time
- t is the irradiation time
- ϕ is the flux within the reactor
- σ_c is the radiative capture cross-section
- σ_f is the fission cross-section
- λ is the decay constant

With a half-life of 87.7 years the decay constant for ^{238}Pu is not negligible. Like the decay of ^{241}Pu to ^{241}Am , the decay of ^{238}Pu to ^{234}U can be used to determine temporal information related to the fuel sample. Depending on the use of the information, it may be important to back calculate to discharge isotopics or use the specific information to compare against available capabilities to identify or eliminate locations of production. Also, the production and removal terms for the intermediate steps in the production timeline do not include all loss scenarios. The ^{238}Pu production also can be used to

determine if the reactor fuel would have been recycled, as recycled fuel has a higher initial ^{236}U content, which would drive a higher ^{238}Pu production. If you are interested in the specific production and loss mechanisms, see the following reference. (Jordheim D., 1990)

Fuel and Reactor Conditions Impacting Pu Production

URANIUM ENRICHMENT

The initial uranium enrichment and isotopics will determine the total burnup achievable by the fuel. This maximum burnup will depend on core configurations, power profiles, and functionality of the reactor; and the fuel may also be discharged early for other reasons unrelated to fuel performance. The amount of ^{235}U will also influence ^{236}U production and eventually ^{238}Pu production, which will affect the isotopic ratios of $^{238}\text{Pu}/^{239}\text{Pu}$. The ^{236}U effect will also be seen in recycled uranium fuel, as the initial isotopics will have ^{236}U present. Refer to Table 1 for the relevant cross section information and natural abundance of uranium.

While the development of the classification tool focused on the plutonium isotopes, there is much information to gain from categorizing, measuring, or calculating the uranium isotopics for inclusion in the classification model and developed tool. The purpose of this method was to test the applicability of a multivariate approach to a specific dataset. If complete datasets are available with more isotopes, it would be the next logical step to incorporate them into this method.

MODERATOR TYPE AND TEMPERATURE

The type of moderator that is used in the reactor design will have an impact on final plutonium isotopics due to a change in energy spectrum. The cross-section figures show

the large resonances for some of the plutonium and uranium isotopes of interest. This change can be seen by changing both the moderator type and the moderator temperature.

As the temperature increases for a reactor, the total system temperature changes which affects all materials in the system, ultimately altering the cross sections. For the fuel, moderator, cladding materials, etc., Doppler broadening occurs in the cross sections, with a large impact in the resonance region of the fuel. This will decrease the resonance escape probability for the neutron giving it a higher likelihood that it will be absorbed and move up the plutonium production chain. These temperature differences can be exploited to determine the operating temperature of the region the material was produced.

Figure 14 shows the elastic scattering and absorption cross sections for the three moderator types of interest: light water, heavy water, and graphite. Things to note in the graph include the high scattering cross section with respect to the absorption cross sections. However, it is important to note for the light water moderator that the absorption cross section is 2-3 orders of magnitude higher than the heavy water and graphite moderators. This larger cross section requires the use of enriched fuel to maintain a critical system.

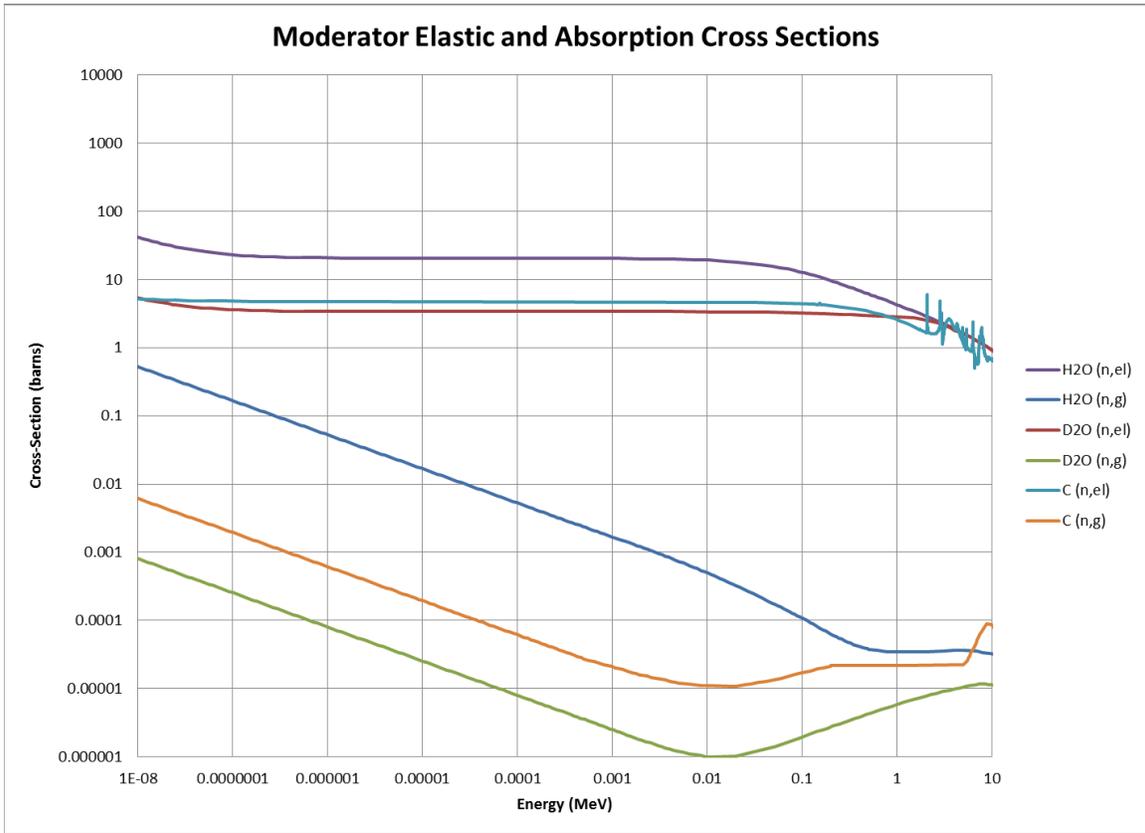


Figure 14: Moderator Elastic and Absorption Cross Sections (Chadwick M., 2011)

REACTOR POWER LEVEL

The reactor power level is a measure of the fission rate in the reactor core. With one fission releasing on average 200 MeV, a 1 MW reactor has a fission rate of approximately 3.12×10^{16} fissions per second. While the power level does not directly affect the plutonium isotopics, it will impact the overall temperature of the core as discussed under Moderator Type and Temperature and will also have an impact on the rate of production and the amount of time a specific fuel type can be irradiated. The power level of a reactor is one of the parameters in calculating the fuel burnup along with irradiation time and fuel mass. Burnup is computed by multiplying the thermal power of the plant by the time of operation and dividing by the mass of the initial fuel loading. For power

reactors, this value is often expressed in terms of MW-d/MTHM (MegaWatt-days per Metric Ton of Heavy Metal), where MTHM includes the initial loading of actinides (uranium, plutonium, etc.).

IRRADIATION TIME

The amount of time the fuel or a target has undergone irradiation at the equivalent of full reactor power is the irradiation time. This value accounts for time the reactor operates at a lower power level or is shut down completely with a parameter known as capacity factor. If the reactor operates at full power every day for a month, it is determined that the reactor was operating at 100% capacity factor for the month. This value will go down if the reactor operates at lower power as it is not utilizing its full capacity. When calculating the burnup of a reactor, the term effective full power days (EFPD) is commonly used to account for the capacity factor of the reactor.

VOID COEFFICIENT FOR BOILING WATER REACTORS

Fuel characterization in boiling water reactors is difficult due to the dependence of the axial location of the fuel. The void coefficient is an important parameter for these reactor types as the fuel temperature, moderator temperature, and neutron spectrum are different in the steam space compared to the liquid region of the core.

CLASSIFICATION MODEL DEVELOPMENT

The previous section discussed the parameters and gave general information on the production of plutonium in a reactor. The knowledge of reactor operation and reactor physics can be exploited in terms of plutonium production. With the understanding of how moderator type, temperature, and burnup can affect the system, correlations to a specific reactor type can be determined based on isotopic data measurements of the finished plutonium product.

The nuclear forensics community has interest in the determination of the origin of interdicted nuclear material. By taking measured data and filling in gaps with calculated data, an information database can be developed that feeds tools for classification. By combining statistical analysis tools with the measured and calculated data, a classification algorithm can be developed that can determine the probability that a plutonium sample came from a specific reactor type.

The initial work in the development of the classification tool used two different approaches to attack the problem. These efforts were worked on concurrently, with one approach focusing on plutonium signatures to determine if reactor specific information could be identified by linear regression of the specific isotope ratios. The other approach was to collect measured data and identify trends in the plutonium ratios that could be exploited for classification. Ultimately these two approaches were combined and further evolved using multivariate regression to build a classification tool that can be used to give a probability of a reactor type of origin. The two approaches will be separated out to capture the work completed on these early tasks before they were ultimately combined to develop the final tool.

SFCOMPO Data Collection and Plutonium Signatures Analysis

SFCOMPO - COLLECTION OF AVAILABLE MEASUREMENT DATA AND PRELIMINARY CHARACTERIZATION

As stated in the background section the measured data from the Spent Fuel COMPOSITION database was used as the starting point to baseline the SFCOMPO/NF database. This primarily supported the approach of data collection and analysis to determine if there were characteristics in the data that could be exploited for classification. Starting with the original database, the following reports provided additional data to increase the range of reactor types available for comparison.

PWR – TMI, Turkey Point and Gösgen

- SCALE 5.1 Predictions of PWR Spent Nuclear Fuel Isotopic Compositions, Oak Ridge National Laboratory (Radulescu G., 2010)

PWR – TMI and BWR – Quad Cities

- Analysis of Spent Fuel Nuclear Samples from Three Mile Island and the Quad Cities Reactors: Final Report, Argonne National Laboratory (Wolf S., 2000)

BWR – Fukushima-Daini2

- Compilation of Measurement and Analysis Results of Isotopic Inventories of Spent BWR Fuels (Yamamoto T., 2009)

Graphite Reactor – B Reactor, K West Reactor and C Reactor

- Experimental Isotopic Analysis of Point Exposure Data in Hanford Production Reactor Fuels, Douglas United Nuclear Inc. (Toffer H., 1970)

CANDU – NPD, Bruce A and Pickering

- Verification and Validation of the ORIGEN-S Code and Nuclear Data Libraries, Atomic Energy of Canada Limited. (Gauld I., 1995)

MAGNOX – Hunterston A, Bradwell

- FISPIN10 Validation Review, British Nuclear Fuels. (Parker D., 2001)

VVER – Novovoronezh NPP-4

- Radiochemical Assays of Irradiated VVER-440 Fuel for Use in Spent Fuel Burnup Credit Activities, Lawrence Livermore National Laboratory. (Jardine L., 2005)

VVER – Novovoronezh NPP-3, NPP-4, NPP-5, Kalinin, Balakovo and RBMK-Leningrad

- Destructive Analysis of the Nuclide Composition of Spent Fuel of VVER-440, VVER-1000 and RBMK-1000 Reactors, Khlopin Radium Institute. (Makarova T., 2008)

This effort created a combined SFCOMPO/NF database that was used for development of the classification algorithm for nuclear forensics (NF) purposes. Table 2 shows the range of operating parameters in the version of the SFCOMPO/NF database by reactor type.

Reactor Type	Enrichment (wt% ²³⁵ U)	Burnup (GWd/MTU)	Cooling time* (years)
PWR	2.453 - 5.07	0.00891 – 55.7	0 – 10.8
BWR	0.71 – 4.5	2.21 – 59.1	0 – 6.7
Graphite PR	0.143 – 2.1	0.259 – 2.083	0.107 – 0.249
CANDU	0.71	6.2	0.499 – 15.315
MAGNOX	0.71	3.90 - 8.93	0.704 – 1.74
VVER	3.3 – 4.4	8.7 – 51.7	3.0 – 10.4
RBMK	1.8 – 2.09	6.2 – 27.7	0.80 – 4.7

*The cooling times of 0 listed above have likely been mathematically altered to obtain this value. Samples are not typically analyzed within seconds or even hours of removal from a reactor. Additionally, cooling times of multiple years will have an impact on the ²⁴¹Pu as it decays to ²⁴¹Am and results in a value lower than would be seen at fuel discharge.

Table 2: Range of Operating Parameters in Current SFCOMPO/NF by Reactor Type

This method will rely on the use of plutonium isotopic ratios as a basis for comparison. These ratios can be plotted against each other in multiple dimensions to

determine a probability for classification. Figure 15, Figure 16, and Figure 17 show measured data from the SFCOMPO/NF database for multiple reactor types for the ratios of $^{238}\text{Pu}/^{239}\text{Pu}$, $^{241}\text{Pu}/^{239}\text{Pu}$, and $^{242}\text{Pu}/^{239}\text{Pu}$. Note that there are general trends in the data which were hypothesized could be used to determine reactor origin. Note that not all the measured data has been decay corrected back to reactor discharge, which will have a noticeable impact on the ^{241}Pu and the $^{241}\text{Pu}/^{239}\text{Pu}$ ratio. The data used from the SFCOMPO/NF database contained only the ^{241}Pu data and did not contain any data on ^{241}Am that is produced as a decay product. The effect of not correcting for ^{241}Pu decay can be seen in Figure 16 with the PWR, BWR, and CANDU data showing a flattening of the $^{241}\text{Pu}/^{239}\text{Pu}$ ratio as burnup increases. For the purposes on this dissertation, the decision was made to include the data “as is” without applying a decay correction as plutonium samples that are analyzed using the classification tool may not be age corrected to discharge isotopics. Additionally, having a large range of different isotopic ratios that encompass the varying landscape of operational parameters and discharge decay times will make the classification tool more robust in identification. If decay correction is necessary in the future, there several publications to reference for that process, (Sampson T., 1986) is an example.

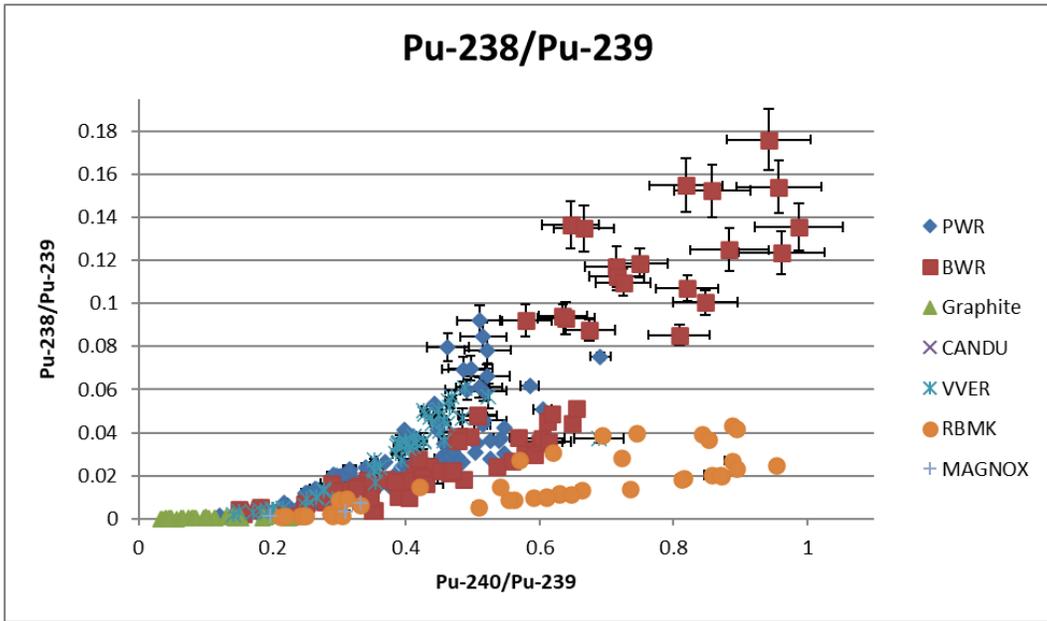


Figure 15: $^{238}\text{Pu}/^{239}\text{Pu}$ Plotted for Multiple Reactor Types

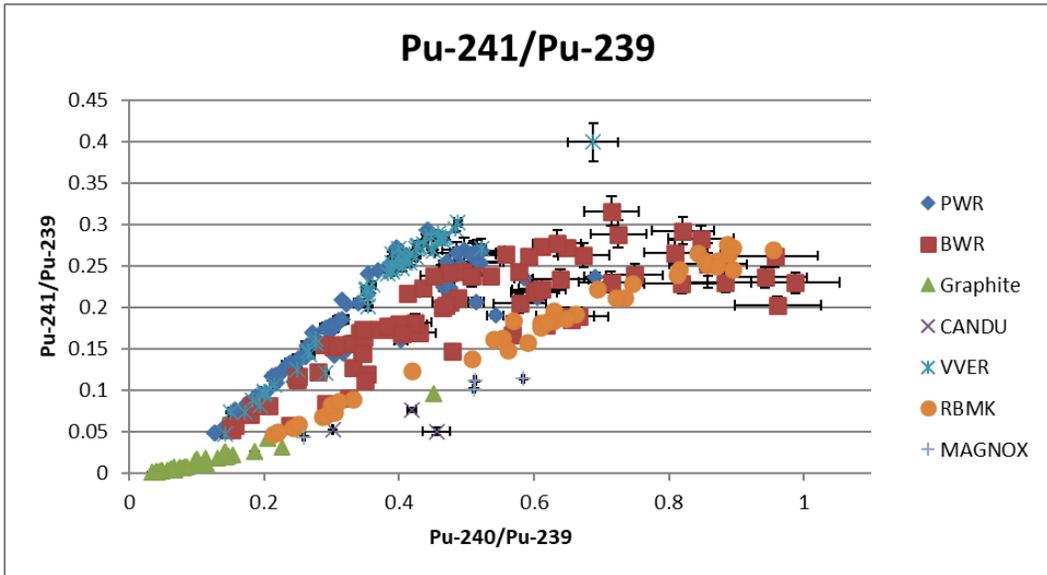


Figure 16: $^{241}\text{Pu}/^{239}\text{Pu}$ Plotted for Multiple Reactor Types

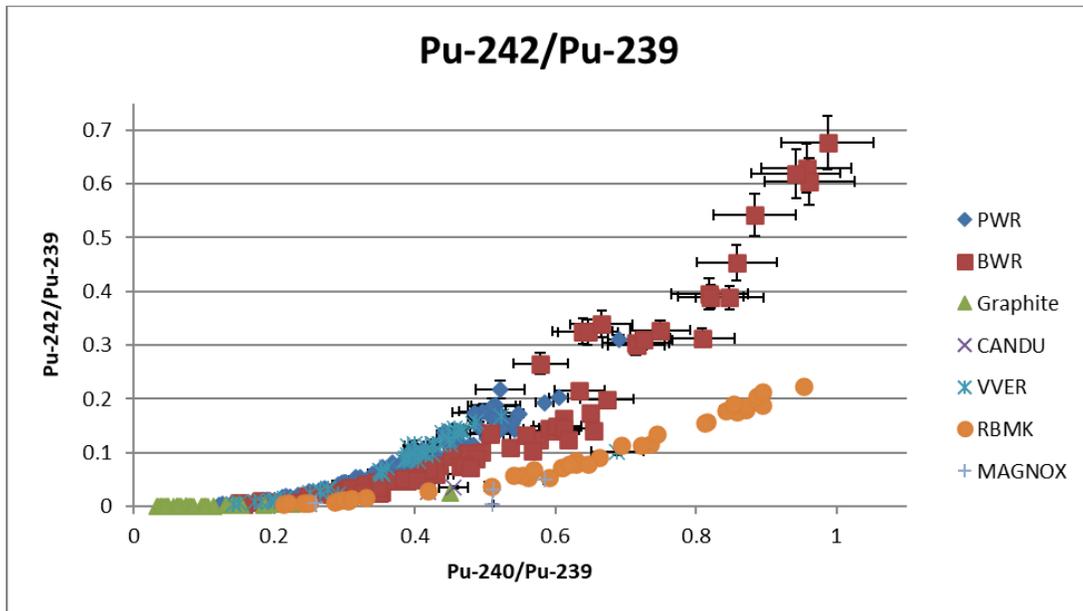


Figure 17: $^{242}\text{Pu}/^{239}\text{Pu}$ Plotted for Multiple Reactor Types

By combining all the plutonium ratios that we can measure and compare against previous material databases, we are able to begin the reactor type classification process of a measured or calculated sample of plutonium.

This is the first step in reactor type classification. The SFCOMPO/NF database contains a large amount of isotopic data, however there are some datasets with a lack of information. Calculated data can be generated using reactor physics models of the different reactor types. Once enough data is calculated to fill in the gaps, the additional reactor types can be added to the reactor classification tool.

This first look at the data identified potential correlations to plutonium isotopic ratios and reactor type of origin. Plutonium ratios of $^{238}\text{Pu}/^{239}\text{Pu}$, $^{240}\text{Pu}/^{239}\text{Pu}$, $^{241}\text{Pu}/^{239}\text{Pu}$, and $^{242}\text{Pu}/^{239}\text{Pu}$ were created to compare data from different reactor types. With Figure 15, Figure 16, and Figure 17 as a basis for the comparison of reactor types, Figure 18, Figure 19, and Figure 20 show each reactor's data fitted with a simple second order least squares

fit with uncertainty bands optimized to encompass $\geq 60\%$ of the data, with the $^{240}\text{Pu}/^{239}\text{Pu}$ as the x-axis for all figures. As shown in the figures, general trends were observed for the different ratios that correspond to different reactor types. This led to an effort to characterize the probability of identification of a reactor type given plutonium isotope ratios.

Note that the data obtained has measurement uncertainties, however, it was not consistently reported in the database, and in some cases, it was omitted. Quantifying these uncertainties is important since there will be inherent uncertainties in measured data used for the classification algorithm, uncertainties in calculated data to fill in data gaps, and uncertainties in measurements of material that is awaiting classification. While the uncertainties reduce the ability to correctly identify a potential reactor type of interest, they also convey the accurate probabilities that a different reactor type is likely.

With all the measured data compiled into a spreadsheet, the ratios of the plutonium isotopes could be compared for the different reactor types. There are five plutonium isotopes of interest that are generated through various decay and neutron capture reactions: ^{238}Pu , ^{239}Pu , ^{240}Pu , ^{241}Pu , and ^{242}Pu . Since ^{239}Pu is the most prevalent, it is standard practice to ratio the other plutonium isotopes to ^{239}Pu . The data comparison highlights the overlap in data giving a probability that a specific ratio comparison can be identified as a specific reactor. For this analysis, only plutonium isotopic ratios were used to compare the different reactor types. Other isotopes including actinides and fission products may allow for better identification of reactor type.

The next step was to identify appropriate statistical methods for characterizing the data, including measurement or calculational uncertainties, where appropriate. This section includes subsections that identify process steps working towards identifying an appropriate classification algorithm using only the calculated data. The knowledge gained during these

steps ultimately led to the final multivariate regression algorithm.

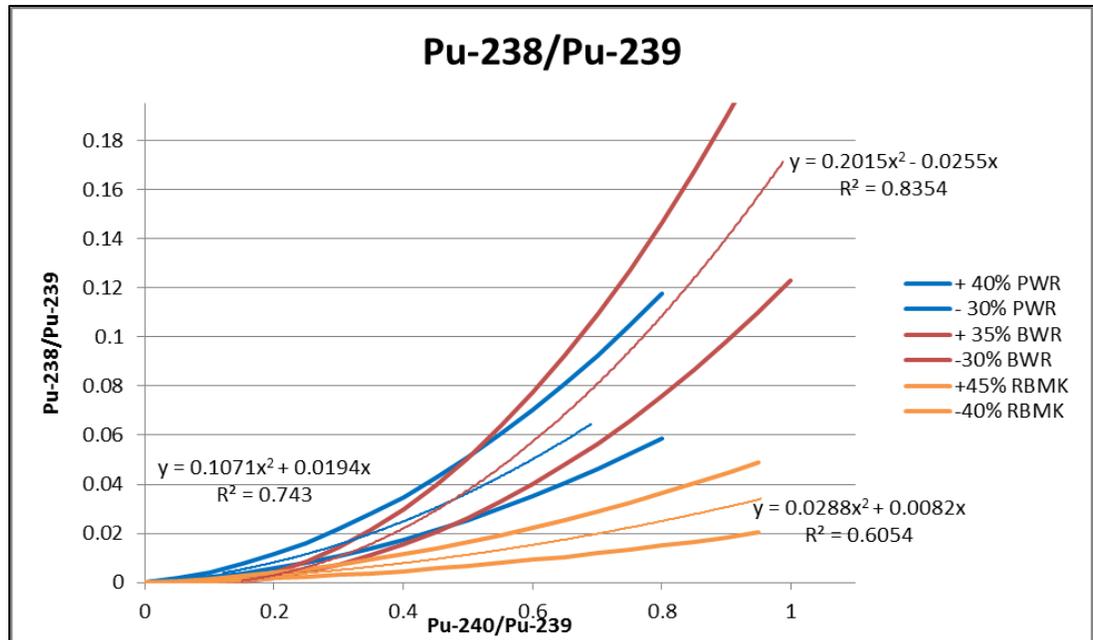


Figure 18: PWR, BWR, and RBMK $^{238}\text{Pu}/^{239}\text{Pu}$ Measurement Data Comparison with Second Order Least Squares Fits and Uncertainty Bands

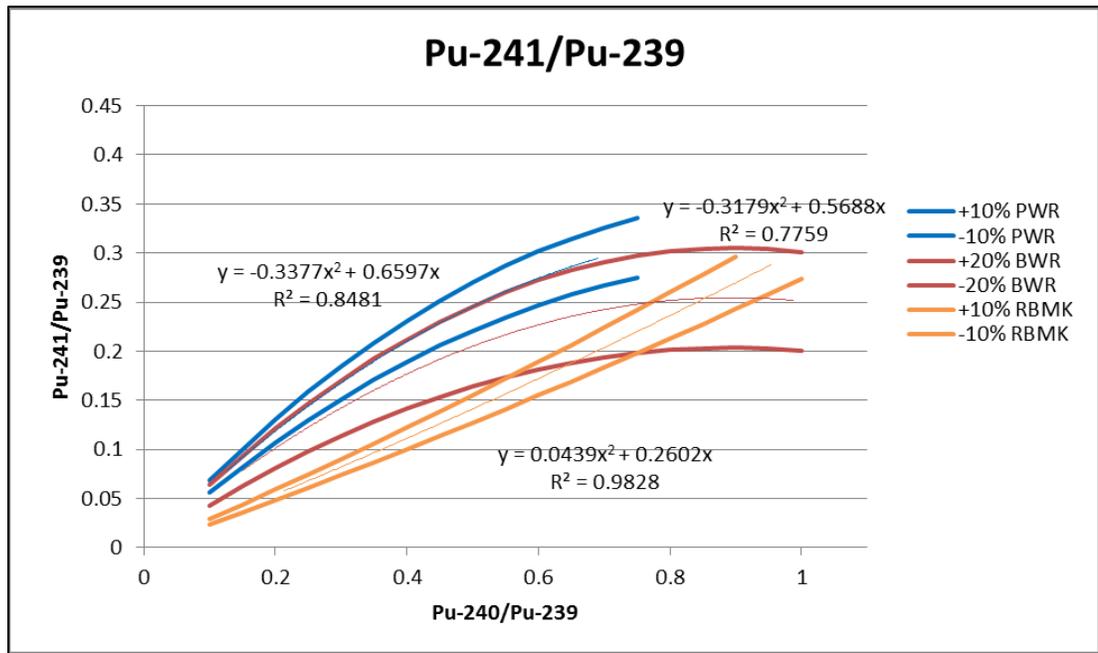


Figure 19: PWR, BWR, and RBMK $^{241}\text{Pu}/^{239}\text{Pu}$ Measurement Data Comparison with Second Order Least Squares Fits and Uncertainty Bands

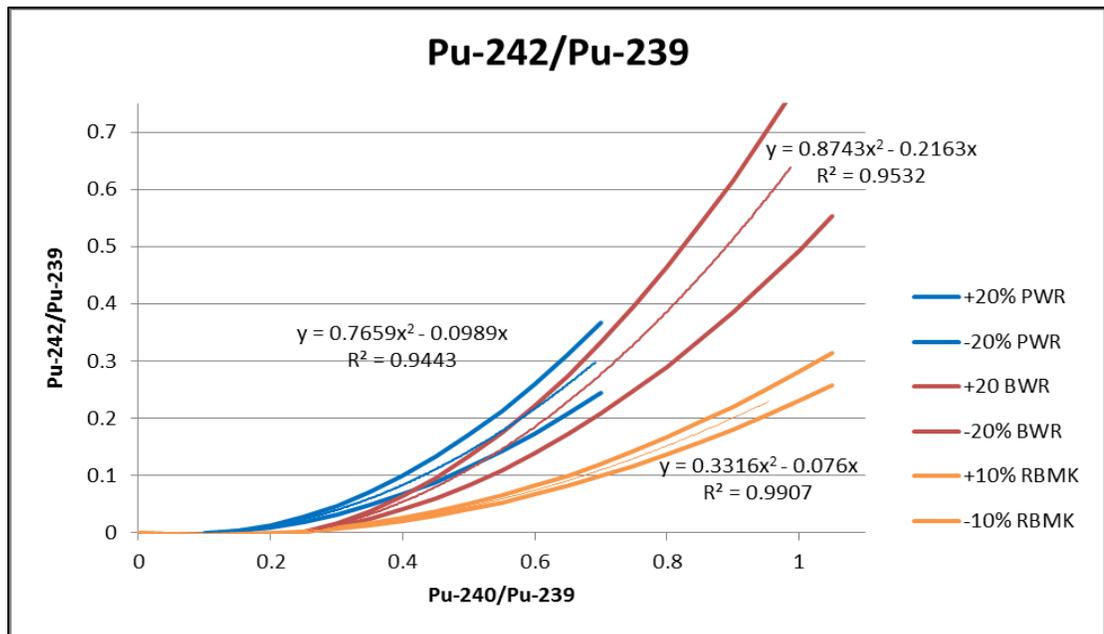


Figure 20: PWR, BWR, and RBMK $^{242}\text{Pu}/^{239}\text{Pu}$ Measurement Data Comparison with Second Order Least Squares Fits and Uncertainty Bands

PU SIGNATURES - GENERATION OF CALCULATED DATA TO EVALUATE ISOTOPIC TRENDS

For the other approach in developing reactor classification capabilities, reactor physics models were created and run to produce “burn-up” curves for each type of reactor, generating the calculated data that was ultimately used for the analysis. Additionally, statisticians were brought on to the project to properly apply appropriate data tools and eventually regression algorithms. Burn-up curves describe the relationship between burn-up (the independent variable), and various isotopic ratios that will be measured. For this evaluation, the four isotopic ratios $^{238}\text{Pu}/^{239}\text{Pu}$, $^{240}\text{Pu}/^{239}\text{Pu}$, $^{241}\text{Pu}/^{239}\text{Pu}$, and $^{242}\text{Pu}/^{239}\text{Pu}$ are used, previously discussed in the background section by Luksic and others. Therefore, the data that is to be used to classify a particular fuel rod consists of a vector $Y = (Y1, Y2, Y3, Y4)$ consisting of the above four ratios. The reactor-physics runs provide us with a “best estimate” of these four isotopic ratios at each burn-up. This “best estimate” will be represented with the term $\mu(B)$; where μ is a vector consisting of the 4 isotopic ratios, and this vector is a function of the burnup, represented by B . The term $\mu(B)$ represents a curve in 4 dimensions relating the isotopic ratios to burnup. If there were no uncertainty in the reactor-physics calculation or in the isotopic measurements, we would expect $Y = \mu(B)$ for some burn-up.

So, if there are four categories of reactor, one would run out four burn-up curves, as identified by the notation $\mu_i(B)$ with $i = 1, 2, 3, 4$ signifying the 4 categories. The measurement Y would be placed in category i only if $Y = \mu_i(B)$ for some burnup B . If the four curves do not intersect, this produces a unique classification scheme. Figure 21 shows the different Pu ratios for the four calculated datasets including a CANDU, Westinghouse 17x17, GE 9x9, and VVER-1000.

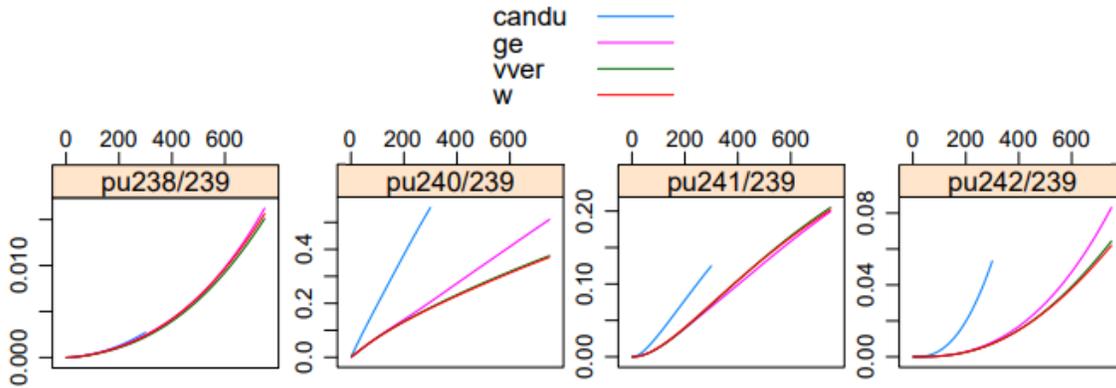


Figure 21: Pu Ratios of Calculated Datasets for Commercial Reactors

The graphs show the differences in isotope production for the different reactor types. There is a noticeable difference between the heavy-water moderated CANDU to the light-water moderated Westinghouse, GE, and VVER-1000 reactors. Additionally, there is a slight difference between the boiling water reactor (BWR) GE, and the pressurized water reactor (PWR) Westinghouse and VVER-1000. These differences are what would allow for statistical analyses to determine a reactor type of origin from a plutonium sample.

PU SIGNATURES – PRELIMINARY ANALYSIS OF CALCULATED DATA

The calculated data generated from ORIGEN-ARP² (ORNL, 2011) included multiple isotopes beyond the plutonium isotopes of interest so analysis to determine if there were isotopic ratio comparisons that would allow for reactor type identification was attempted. Multiple pairwise concentrations progressing through time were generated using the calculated isotopic data for Np, U, Pu, and Am. While some ratio comparisons showed promise, unfortunately the isotopes involved were either unmeasured in actual

² ORIGEN-ARP is a SCALE isotopic depletion and decay analysis sequence used to perform point-depletion calculations with the well-known ORIGEN-S code using problem-dependent cross sections. Problem-dependent cross-section libraries are generated using the ARP (Automatic Rapid Processing) module using an interpolation algorithm that operates on pre-generated libraries created for a range of fuel properties and operating conditions.

samples, or the measurements had large uncertainties that were not quantified in the statistical analysis of the calculated data. Additionally, some of these isotopes were calculated at levels much below the minimum detection level.

Figure 22 shows the analysis relating to a regression using the calculated data. With the differences in the CANDU and BWR apparent, and the Westinghouse and VVER Reactors both being PWR type reactors, there is enough difference in the data that a regression can identify the reactor type with 100% probability.

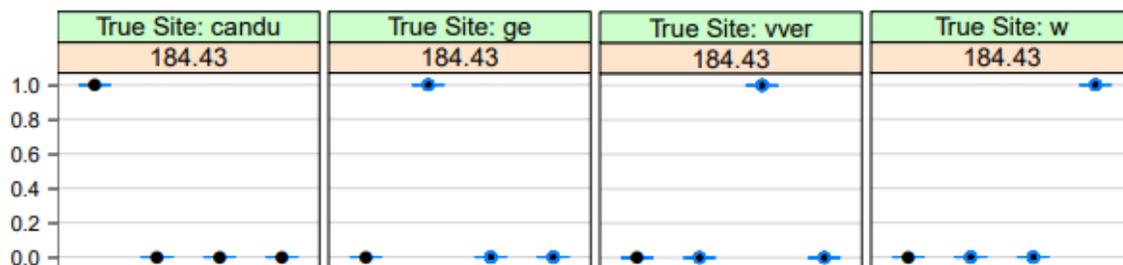


Figure 22: Posterior probability that sample is from tested site given that it came from true site for all four reactor types at a given timestep

Unfortunately, measured data has overlap between the two different reactor types and uncertainty in both measured and calculated data would result in a lower probability of classification. To evaluate the impact of uncertainty in the data, a broad uncertainty was applied to the calculated data to determine if reactor type identification was still possible. Uncertainty in the data, both measured and calculated, will be discussed how it is applied in the different sections and in the overall classification tool.

PU SIGNATURES – ADDITION OF UNCERTAINTY TO THE REGRESSION MODEL

To test the statistical analysis method developed using calculated data, uncertainty was added to the data. For the purposes of this test, a blanket 3% uncertainty was applied

to all the calculated data. This analysis demonstrated that including uncertainties in the data makes it more difficult to discriminate amongst the various types of PWRs. Since it was determined that the CANDU data and the LWR data were sufficiently different, the CANDU data was removed from the plot to highlight if there were significant differences in the LWR data. Figure 23 shows plots of $^{238}\text{Pu}/^{239}\text{Pu}$ compared to $^{240}\text{Pu}/^{239}\text{Pu}$ for the GE, Westinghouse and VVER-1000 calculated data with the added 3% uncertainty.

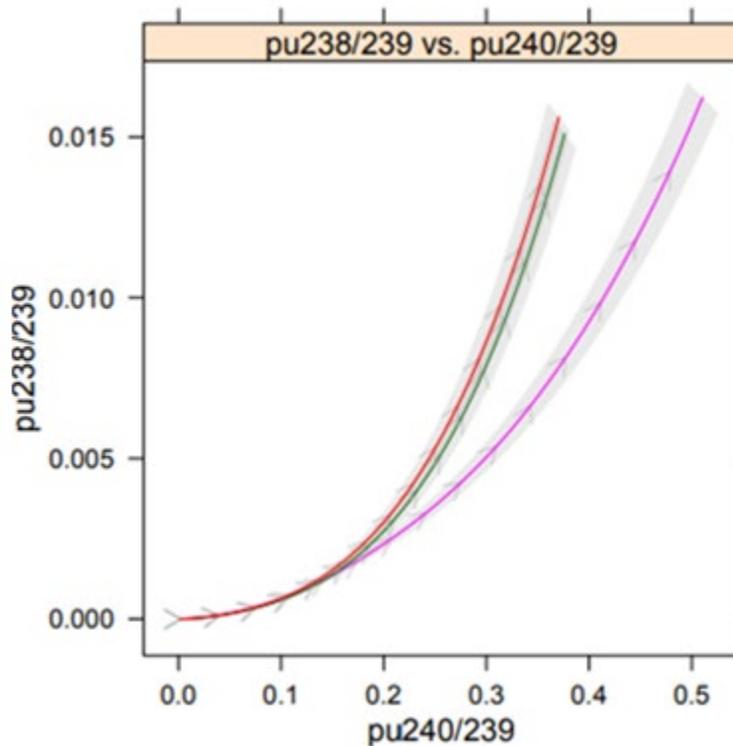


Figure 23: Comparison of $^{238}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ for the Calculated Dataset

It should be noted that this original formulation did not propose any methodology for determining how the uncertainty, as represented by the covariance matrices C_i should be determined. This is an important parameter, and if the uncertainty is not correctly

quantified, the resulting algorithm will work poorly, which is documented in the subsequent subsections.

PU SIGNATURES – GENERATING A REGRESSION MODEL FROM THE CALCULATED DATA

With the data generated and characterized into the pairwise plots, a linear regression model was applied to the data to determine the probability of correct characterization. The data used the uncertainty of three percent for each isotope. As stated before, the isotopic differences from the CANDU to the three LWRs were easily distinguished in the model. Additionally, the initial results performed very well in identifying the correct reactor type based on the data input to the model at the different timesteps of reactor operation. Figure 24 shows the posterior probabilities for the four reactor types calculated data with data uncertainty of 3%. Note that the model correctly identifies the reactor type of origin for the CANDU and BWR (GE), but we start to see posterior probabilities less than one for the VVER and Westinghouse data, which is expected since both are the same PWR type.

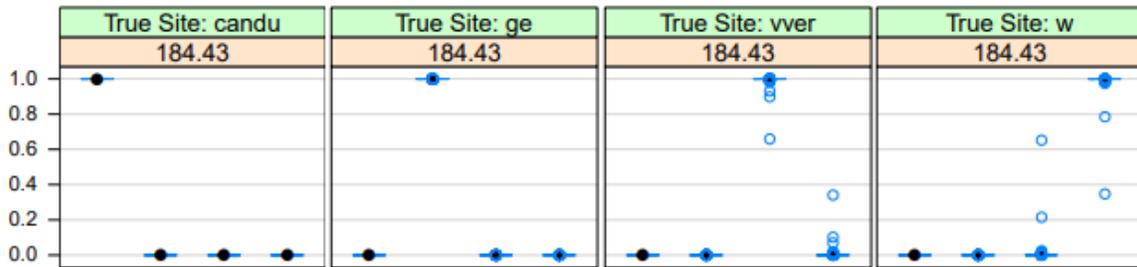


Figure 24: Posterior probability that sample is from tested site given that it came from true site for all four reactor types with 3% uncertainty at a given timestep

Since the model performed very well using the 3 percent isotope uncertainty, the next step in analysis was to use a more conservative uncertainty of 15 percent to determine if classification was still possible. The CANDU reactor was still easily identifiable because

of the difference in spectrum that resulted in noticeably different isotopic ratios. Additionally, the BWR (GE) reactor had posterior probabilities of about 0.8. The two PWR (VVER and Westinghouse) reactor types had posterior probabilities less 0.6. This is expected since all three reactors are LWRs and the VVER and Westinghouse are both PWRs. This analysis indicated that the correct identification of data uncertainty is important for correct reactor type identification. Figure 25 shows the posterior probabilities with a data uncertainty of 15%. Note that in this example, the model gave a higher probability of 0.55 to the “VVER”, however, the physics of the two reactor types are similar enough that there was a probably to misclassify the reactor as a Westinghouse design noted by the probability of approximately 0.35, with the remainder going to the BWR (GE).

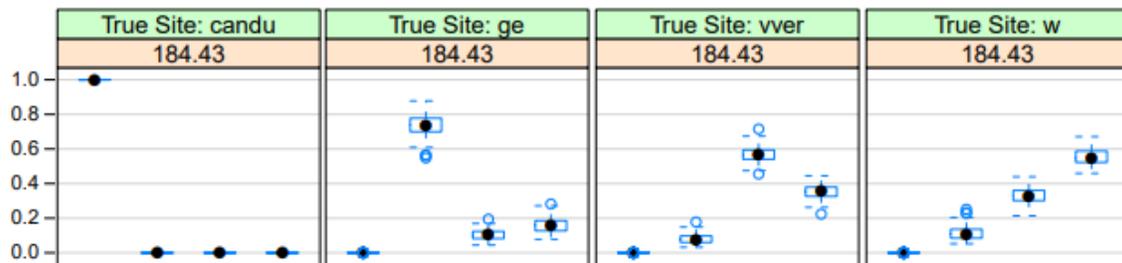


Figure 25: Posterior probability that sample is from tested site given that it came from true site for all four reactor types with 15% uncertainty at a given timestep

PU SIGNATURES – TESTING OF REGRESSION MODEL

Now that the regression model had been developed, the next step was to test the model using external data to determine the performance. A single point of Pu isotopics was generated using a reactor physics code and applied to the model. Table 3 shows the generated data, which was a PWR.

Isotope	Value
²³⁸ Pu	1.04E-04
²³⁹ Pu	6.52E-03
²⁴⁰ Pu	2.15E-03
²⁴¹ Pu	1.67E-03
²⁴² Pu	4.14E-04

Table 3: Pu Isotopics Generated for Regression Model Test

The model was unable to determine the reactor type of origin based on the given isotopics. Figure 26 shows how the generated data compares to the model for the isotopic ratios of interest, with the different Pu ratios compared to the ²⁴⁰Pu/²³⁹Pu. Note that there is no clear reactor type that is identified from this model, however, the data point falls into the area of uncertainty for both the Westinghouse and VVER reactors, of which both are PWRs. Again, this method relied on only calculated data, and did not take advantage of the measured data in the SFCOMPO/NF database. The next step in the approach was to expand on the linear regression into a multivariate regression algorithm utilizing the available data.

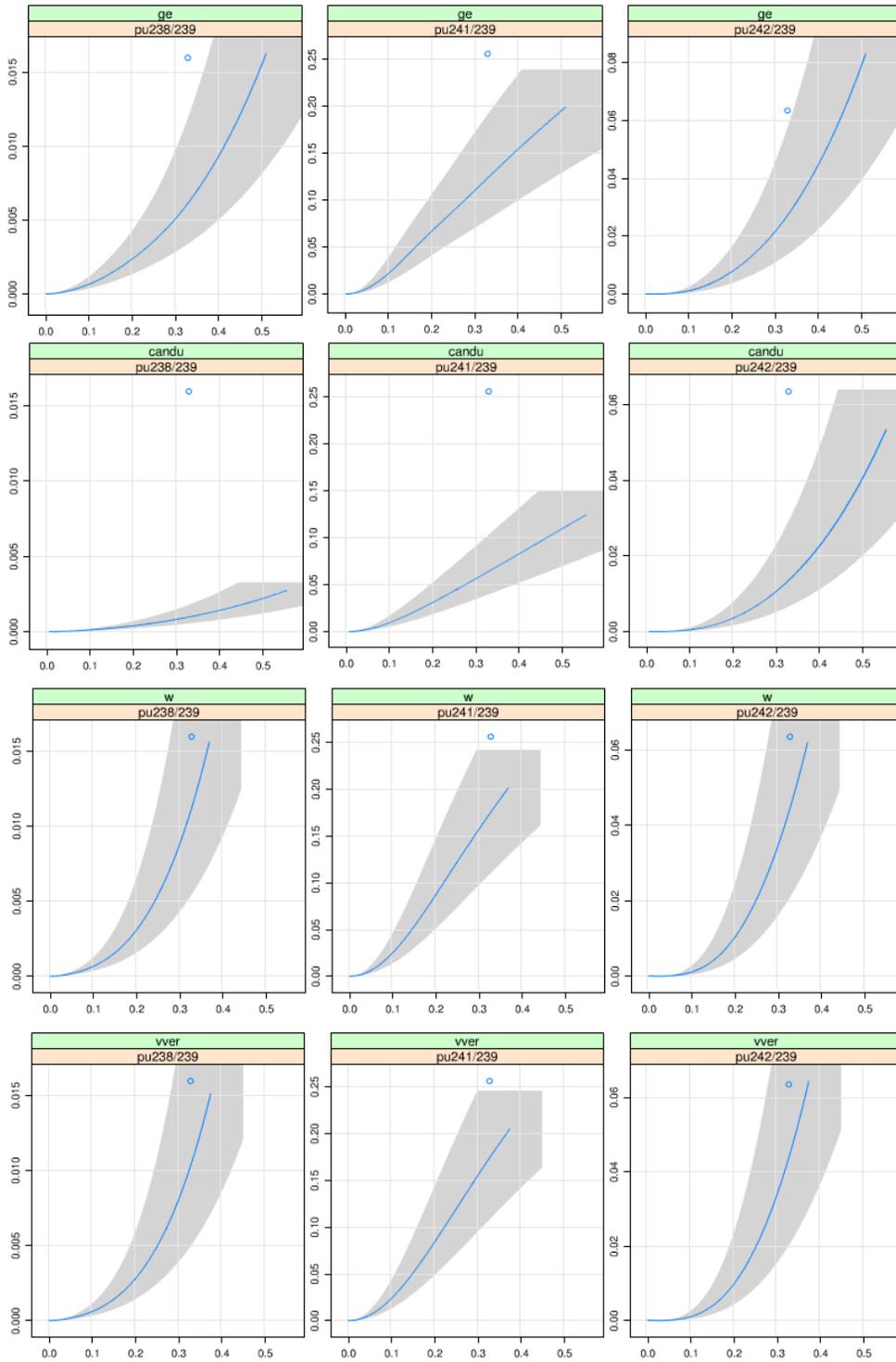


Figure 26: Testing of Regression Model Using Measured Data

Multivariate Characterization of Measured Data (Version 1)

Since it was determined from the analysis of calculated data, a linear regression approach based on burnup was insufficient for reactor origin identification, the next step was to determine if a multivariate approach to statistical analysis would lead to better identification and classification of a reactor type by utilizing all four different plutonium isotopic ratios. An additional statistician with more experience in isotopic ratios was employed to help apply a multivariate approach to the measured and calculated data. The first version of the confusion matrix utilized data from the SFCOMPO/NF database and focused on the different reactor types that contained enough data points for classification. The classification of interest was reactor type, which originally consisted of the categories:

- GE: Boiling Water Reactor designed by GE, (LWR)
- Hanford: Single Pass Graphite Pu production reactors (excludes N-reactor) at Hanford
- RBMK-1000: Russian RBMK series reactor
- Westinghouse: Pressurized Water Reactor designed by Westinghouse, (LWR)
- VVER-1000: Pressurized Water Reactor designed by Russia, (LWR)
- VVER-440: Pressurized Water Reactor designed by Russia, (LWR)

Two additional categories present in the SFCOMPO/NF database were not included in the original classification algorithm because of the limited amount of data available. These categories had less than ten data points in the measured data:

- CANDU: Canadian heavy water reactor.
- MAGNOX: British Gas Cooled graphite reactor.

AVAILABLE TRAINING DATA AND REGRESSION MODEL FIT

Because of the available data, the first version of the algorithm was limited to the above six datasets. This methodology continued to evolve in the additional tasks discussed later in this document. Table 4 identifies the number of data points that contained plutonium ratios and the number of data points that contained all four ratios of interest.

	Total	Complete
GE	78	78
Hanford	46	33
RBMK-1000	41	41
Westinghouse	156	118
VVER-1000	14	14
VVER-440	41	40

Table 4: Number of Data Points Containing Pu Ratios Compared to Complete Ratio Sets

It was determined that future measurements would contain all the required plutonium measurements for a complete dataset, so it was not worth generating a more complicated algorithm to account for missing data. As noted in the previous section, uncertainty data was available for some of the measurements, however since not all data contained the information, the code assumed a default value of 0% uncertainty for all data when calculating the covariance matrices.

With the dataset identified, the vector Y, represented by

$$Y = (Y1, Y2, Y3, Y4) = \left(\frac{{}^{238}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{240}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{241}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{242}\text{Pu}}{{}^{239}\text{Pu}} \right)$$

was assembled from approximately 400 fuel rods, originating from 30 different reactors as “training” data. This training data was used to estimate the parameters appearing in the

classification algorithm which will be discussed in further detail in the next section. Since no associated burn-up information was available for each isotopic measurement, it was not possible to produce an empirical version of the reactor-physics burn-up curves, as described by the term $\mu_i(B)$. This was solved by transforming the input data. The data transformation was also chosen to make the uncertainty distribution most closely resemble a multivariate normal.

Burn-up was replaced by the isotopic ratio $^{240}\text{Pu}/^{239}\text{Pu}$ because this ratio is almost linearly related to burnup, B. From plots of the data, it was clear that uncertainty in the isotopic ratios was approximately proportional to magnitude, which indicated a log transform of the data would produce data that fit a multivariate normal. The Transformed Burnup, B' is given by:

$$B' = \log(Y2) = \log(X) = \log\left(\frac{^{240}\text{Pu}}{^{239}\text{Pu}}\right)$$

and the Transformed Isotopic Ratios are, Y', are given by:

$$Y' = \log(Y1, Y3, Y4) = \log\left(\frac{^{238}\text{Pu}}{^{239}\text{Pu}}, \frac{^{241}\text{Pu}}{^{239}\text{Pu}}, \frac{^{242}\text{Pu}}{^{239}\text{Pu}}\right)$$

The burn-up curves now describe the relationship between $\log(^{240}\text{Pu}/^{239}\text{Pu})$ and the other three isotopic ratios. This transformed data is fit to a multivariate regression model of form;

$$Y'_{ij} = \alpha_{1i} + \alpha_{2i}B'_{ij} + E_{ij}$$

where the index $i = 1, 2, \dots, 6$ represents the category of reactor the data originated from and j indexes the measurements in the category. The α represent the unknown regression

parameters to be estimated and determine the shape of the burn-up curve for category i . The regression residuals, E_{ij} determine the covariance matrices using the relationship $Cov(E_{ij}) = C_i$.

A regression fit to the training data will produce the required covariance matrices C_i and the burn-up curve, $\mu_i(B')$, through the formula:

$$\mu_i(B') = \alpha_{1i} + \alpha_{2i}B'$$

From the linear regression approach, it was determined that uncertainty must be considered for the proper classification. Uncertainty in the measurements and calculations means that Y will never lie exactly on the burnup curve $\mu_i(B)$, but only close to it. This shows that uncertainty transforms this into a probabilistic problem. Let C_i represent a covariance matrix that represents the (measurement and calculational) uncertainty that exists between Y (a measurement from category i) and $\mu_i(B)$. Assuming that the measurement Y is multivariate normal, we have;

$$p(Y|i) = \Phi(Y|\mu_i(B), C_i)$$

with $\Phi(Y|\mu, C)$ representing the multivariate-normal density with a mean of μ and covariance of C . With this formulation, the classification problem has a text-book solution; so, a Quadratic Discriminant Algorithm (QDA) or its Bayesian analog, which produces a posterior classification probability vector was applied. The posterior vector, $p(Y) = (p(i = 1|Y), p(i = 2|Y), \dots, p(i = 4|Y))$ has the definition

$$p(Y|i) = H_0 \Phi(Y|\mu_i(B), C_i) \text{prior}_i$$

where *prior* represents the prior vector for the categories; To perform ordinary (non-Bayesian) QDA, the “uniform” *prior* = (1/4, 1/4, ...1/4) is used. The term H_0 represents a normalization constant chosen to force the posterior probabilities to sum to one. Consequently, H_0 must be:

$$H_0 = \left[\sum_{i=1}^6 \Phi(Y|\mu_i(B), C_i) \text{prior}_i \right]^{-1}$$

The above discussion has ignored a difficulty in this formulation; The burnup, B is not known. To circumvent this problem, a regression estimate for burnup is produced, which is called \hat{B}_i and is substituted into the above formulas.

Figure 27, Figure 28, and Figure 29 present the transformed data Y' by the transformed burnup B' for each reactor type, with the regression fits also included on the plots. These plots are in log scale so that the regression fits are linear and show that the log-linear model appears to be generally reasonable. Note that the reactors with similar moderators follow along a similar regression fit.

For reactor classification to be effective, the data for each reactor type should have separation from the other reactor types. However, these plots indicate that this is not the case as most reactor types overlap in data. This indicates that the algorithm might have difficulty in classification. It is important to note that burnup (as defined as $B' = {}^{240}\text{Pu}/{}^{239}\text{Pu}$) has a different range for the different reactor types. Note that the lower burnup Hanford reactors, with different operating conditions and parameters, does not intersect the higher burnup, power producing reactors.

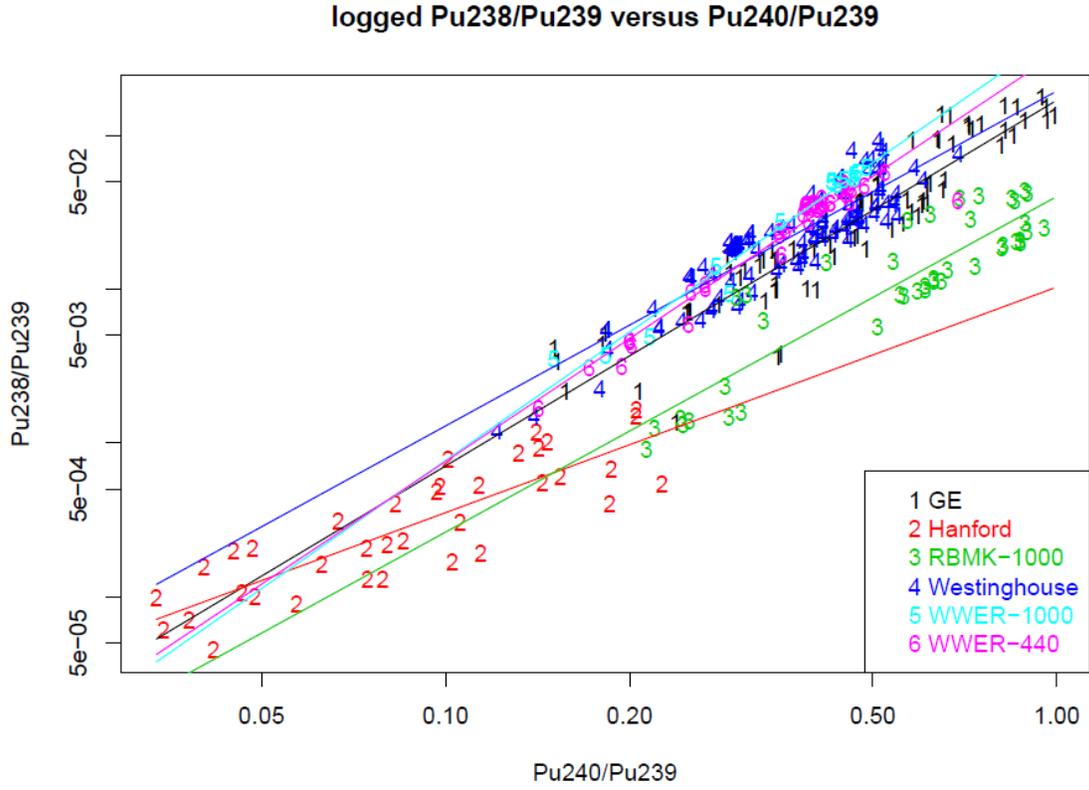


Figure 27: Regression of Different Reactor Types based on $^{238}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ Comparison

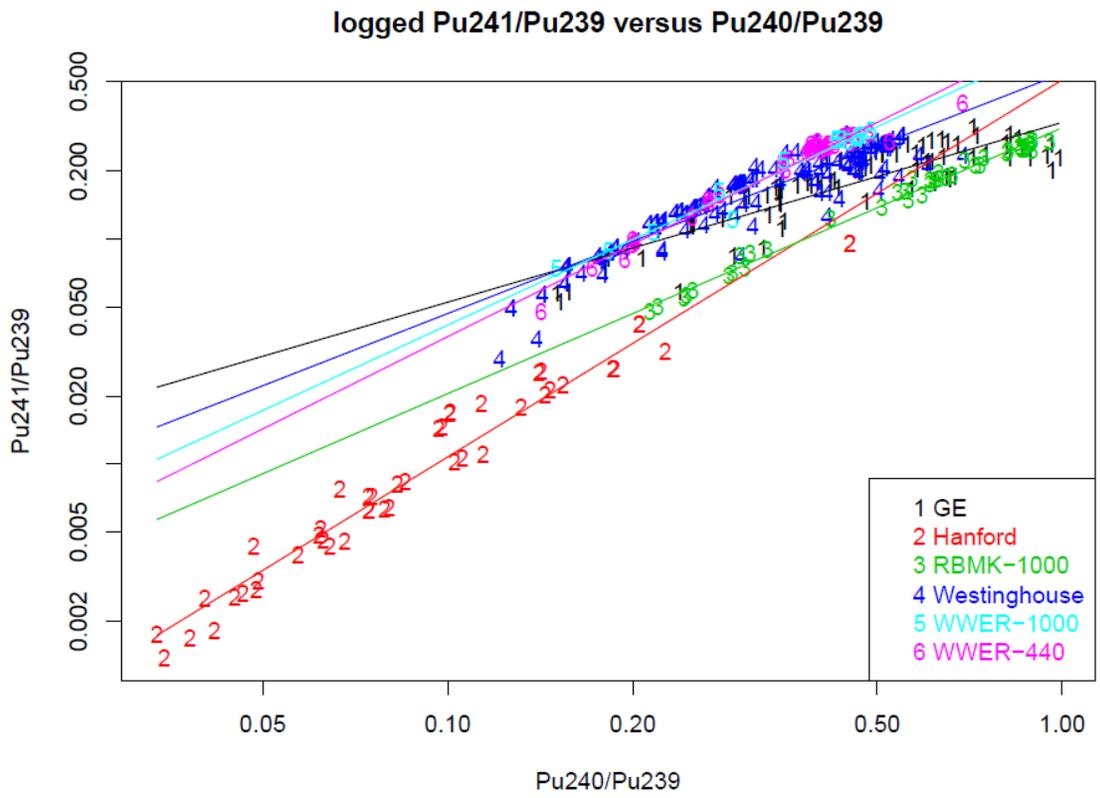


Figure 28: Regression of Different Reactor Types based on $^{241}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ Comparison

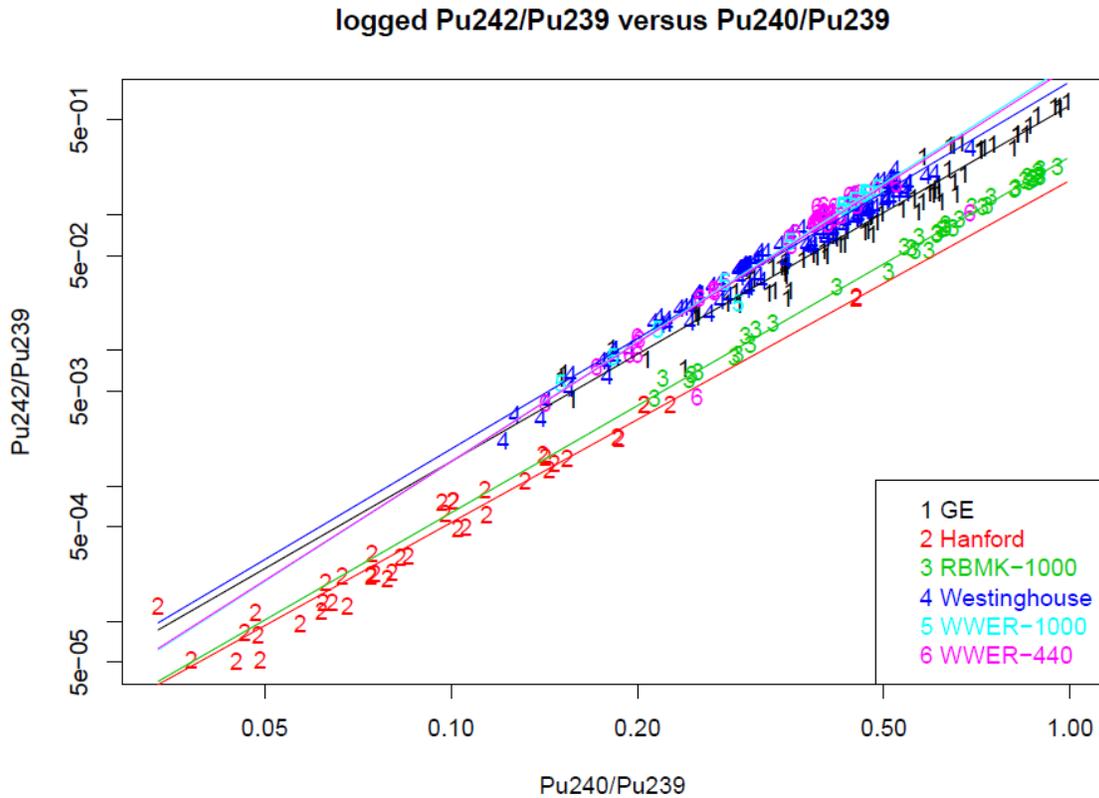


Figure 29: Regression of Different Reactor Types based on $^{242}\text{Pu}/^{239}\text{Pu}$ to $^{240}\text{Pu}/^{239}\text{Pu}$ Comparison

The regression fit to the data produces estimates for the parameters α_{1i} , α_{2i} and C_i .

These parameters estimates are shown in Table 5 and Table 6 and when plugged into

$$p(Y|i) = H_0 \Phi(Y|\mu_i(B), C_i) \text{prior}_i$$

produce a classification algorithm that can be applied to new data. These tables were used to develop an excel workbook to run the classification algorithm against a set of plutonium ratios. These results will be shown in a later section.

j Isotopic Ratio	α_{1ij}	α_{2ij}
	GE (i=1)	
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.777	2.379
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.109	0.799
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-0.445	2.626
	Hanford (i=2)	
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-4.574	1.467
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-0.673	1.676
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-1.717	2.525
	RBMK-1000 (i=3)	
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-3.218	2.183
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.163	1.182
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-1.317	2.627
	Westinghouse (i=4)	
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.639	2.178
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-0.609	1.067
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-0.043	2.709
	VVER-1000 (i=5)	
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-0.79	2.771
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-0.279	1.262
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	0.26	2.934
	VVER-440 (i=6)	
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.076	2.654
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-0.166	1.362
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	0.209	2.913

Table 5: Alpha Parameter Estimates from Regression Fit

	$^{238}\text{Pu}/^{239}\text{Pu}$	$^{241}\text{Pu}/^{239}\text{Pu}$	$^{242}\text{Pu}/^{239}\text{Pu}$
	GE (i=1)		
$^{238}\text{Pu}/^{239}\text{Pu}$	0.1981	0.0097	0.0792
$^{241}\text{Pu}/^{239}\text{Pu}$	0.0097	0.0418	0.0101
$^{242}\text{Pu}/^{239}\text{Pu}$	0.0792	0.0101	0.0499
	Hanford (i=2)		
$^{238}\text{Pu}/^{239}\text{Pu}$	0.2456	0.0809	0.0739
$^{241}\text{Pu}/^{239}\text{Pu}$	0.0809	0.046	0.0329
$^{242}\text{Pu}/^{239}\text{Pu}$	0.0739	0.0329	0.0983
	RBMK-1000 (i=3)		
$^{238}\text{Pu}/^{239}\text{Pu}$	0.2484	0.0223	0.029
$^{241}\text{Pu}/^{239}\text{Pu}$	0.0223	0.0035	0.0032
$^{242}\text{Pu}/^{239}\text{Pu}$	0.029	0.0032	0.0076
	Westinghouse (i=4)		
$^{238}\text{Pu}/^{239}\text{Pu}$	0.1228	0.0344	0.0498
$^{241}\text{Pu}/^{239}\text{Pu}$	0.0344	0.0335	0.0225
$^{242}\text{Pu}/^{239}\text{Pu}$	0.0498	0.0225	0.0263
	VVER-1000 (i=5)		
$^{238}\text{Pu}/^{239}\text{Pu}$	0.0424	0.0154	0.0238
$^{241}\text{Pu}/^{239}\text{Pu}$	0.0154	0.0085	0.012
$^{242}\text{Pu}/^{239}\text{Pu}$	0.0238	0.012	0.0195
	VVER-440 (i=6)		
$^{238}\text{Pu}/^{239}\text{Pu}$	0.0616	0.0146	0.0724
$^{241}\text{Pu}/^{239}\text{Pu}$	0.0146	0.0074	0.016
$^{242}\text{Pu}/^{239}\text{Pu}$	0.0724	0.016	0.1312

Table 6: Covariance Matrix (C_i) Estimates from Regression Fit

To evaluate the general capability of the classification algorithm, cross validation was used on the available SFCOMPO/NF dataset. To apply cross validation, the data was randomly split into two equal halves, with the first half used as the training set to estimate the algorithm parameters shown in Table 5 and Table 6. The fitted algorithm is applied to

the validation set, which is the second half of the data, to produce classification results. This procedure was repeated 300 times using a different random split of the data.

RESULTS OF MULTIVARIATE ANALYSIS USING ONLY MEASURED DATA

The overall performance of the algorithm is quantified by two matrices. The “confusion” matrix, shown in Table 7, describes the proportion of the validation points that fall into each category, given the true category of each point. To calculate the confusion matrix, the algorithm chooses the category with the highest posterior probability calculated from the cross-validation procedure. The second matrix used to quantify the algorithm performance presents the average posterior for all validation points in each category. This information is presented in Table 8.

	GE	Hanford	RBMK-1000	Westinghouse	VVER-1000	VVER-440
GE	0.768	0.026	0.006	0.098	0.055	0.048
Hanford	0.002	0.998	0.000	0.000	0.000	0.000
RBMK-1000	0.002	0.052	0.946	0.000	0.000	0.000
Westinghouse	0.150	0.003	0.000	0.644	0.150	0.054
VVER-1000	0.065	0.000	0.000	0.276	0.546	0.113
VVER-440	0.032	0.048	0.000	0.208	0.138	0.574

Table 7: Confusion Matrix Developed from SFCOMPO/NF Data

	GE	Hanford	RBMK-1000	Westinghouse	VVER-1000	VVER-440
GE	0.773	0.026	0.005	0.099	0.049	0.049
Hanford	0.002	0.997	0.001	0.000	0.000	0.000
RBMK-1000	0.003	0.071	0.926	0.000	0.000	0.000
Westinghouse	0.187	0.003	0.000	0.595	0.129	0.086
VVER-1000	0.091	0.000	0.000	0.281	0.482	0.146
VVER-440	0.049	0.048	0.000	0.273	0.124	0.506

Table 8: Average Posteriors Developed from SFCOMPO/NF Data

DISCUSSION AND FUTURE IMPROVEMENTS

Table 8 shows that the classification algorithm works well for the GE, Hanford, and RBMK-1000 categories with misclassification below 23%, 1%, and 8%, respectively. For the remaining three categories, Westinghouse, VVER-1000, and VVER-440, the misclassification is between 40-50%. However, it should be noted that all three of those categories are the PWR reactor type and could be collapsed into one overall PWR category.

Additionally, it was determined that a larger training dataset was needed to expand the functionality of the model, including increasing the number of reactor types in the dataset and expanding the burnup range of the existing reactor types. The poor performance of the model with data at low or near-zero burnup was also identified, and future work would be needed to address this issue.

Ultimately, the ability to classify the 4 reactor types with a reasonable amount of certainty even though there was significant overlap in the individual ratio plots was promising for future algorithm development.

Multivariate Characterization Using Measured and Calculated Data (Interim Version of Classification Tool Development)

Building on the results of the first model, the next step was to incorporate calculated data from reactor physics models to increase the amount of training data and to incorporate the reactor types that did not have sufficient measured data available for regression analysis. This effort provided an interim version of the classification tool where the alpha parameters and covariance matrices were generated, but the analysis was not incorporated into the classification tool. As stated in the previous section, while there is uncertainty information for some of the measured data, since not all the calculated or measured data contained uncertainty values, a default value of 0% uncertainty was used for all data for the generation of the covariance matrices for the classification tool. The calculated data was generated using ORIGEN-ARP (ORNL, 2011) and was intended to produce a more complete dataset to produce improved regression model fits. It must be noted that the use of the reactor physics models to generate calculated data is that it provides no information on the variability in the isotopic ratios. The data generated represents core average isotopics at a given burnup value for the different reactor types and does not include any uncertainty information for the value. For the purposes of the classification tool development, the calculated data represented a small fraction of the total dataset, and the weighting factor was not utilized. If the proportion of calculated data increases, the appropriate weighting factor would need to be applied to not bias the results to the calculated data.

Calculated data was supplied for Hanford, Magnox, CANDU, Westinghouse, and GE reactor series, and this allowed all reactor series in the original SFCOMPO/NF dataset

to be included in the regression and classification models. The two series (CANDU and Magnox) that were dropped in the first attempt at multivariate regression because of a limited amount of measure data are now in the analysis. Additionally, the low burn-up information from the reactor physics models, allows classification error to be calculated as a function of burn-up.

REGRESSION MODEL UPDATES

Working with the statistician, the previous regression model assumed a simple linear relationship between isotopic ratios on the log scale, but with the addition of the calculated data from the reactor physics models, this model did not fit the expanded dataset. It was determined that the addition of a quadratic term to the model would better fit the new data:

$$Y'_{ij} = \alpha_{1i} + \alpha_{2i}B'_{ij} + \alpha_{3i}(B'_{ij})^2 + E_{ij}$$

where

$$B' = \log(Y2) = \log(X) = \log\left(\frac{{}^{240}\text{Pu}}{{}^{239}\text{Pu}}\right)$$

and

$$Y' = \log(Y1, Y3, Y4) = \log\left(\frac{{}^{238}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{241}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{242}\text{Pu}}{{}^{239}\text{Pu}}\right)$$

Since the calculated data points contained less variability and no uncertainty information (provided a single burnup curve for a reactor cycle), when compiling the data with the measured points, weighting the data appropriately (incorporating the necessary uncertainties generated from the reactor physics models and overall results) will be necessary when to total amount of calculated data becomes significant as the regression

estimate for the covariance could become biased towards the calculated data. For this evaluation, the amount of measured data is much larger than the calculated data, so the additional calculated data added to the dataset is weighted the same as the measured data. If future work expands the training dataset with a larger percentage of calculated data, proper data weighting may become important. The expected impact of changing the weight of the calculated data in a dataset for a reactor type will become more dispersed, which will probably increase the overlap of the different reactor type datasets and ultimately reduce the accuracy in determining a reactor type of origin.

In further discussions with the project statistician, we addressed the issue pertaining to the error structure for low burnup. In the previous regression and classification model, it was assumed that the error was constant on the log scale (equivalently proportional on an unlogged scale), which leads to logical contradictions for low burn-up. With proportional errors, the classification model handles low burnup appropriately, however this does not account for measurement errors and uncertainties at these lower levels. To obtain more realistic behavior at lower burnup, a variance model with the following form was used for an isotopic ratio measurement, R .

$$V(R) = S_d^2 + (S_p R)^2$$

Where S_p represents the proportional error, or relative standard deviation, while S_d represents the constant “detection level” error. Since logged data is used, the equation that is relevant to the classification model is then:

$$V(\log(R)) = \left(\frac{S_d}{R}\right)^2 + S_p^2$$

For this evaluation, S_d was set to 10^{-4} , which implies that a ratio of magnitude less than S_d cannot be detected. This gives an instrument “detection level” the value of S_d . Since $R = \exp(\mu)$, the equations simplify to:

$$V(\mu) = \left(\frac{S_d}{\exp(\mu)}\right)^2 + S_p^2$$

and

$$S(\mu)^2 = \exp(2(\log(S_d) - \mu)) + S_p^2$$

Next, we let the original covariance matrix be written as C_p , and the new covariance matrix associated with the value μ is given by:

$$C(\mu) = \text{Diag}\left(\frac{S(\mu)}{S_p}\right) C_p \text{Diag}\left(\frac{S(\mu)}{S_p}\right)$$

Where the notation $\text{Diag}(x)$ represents a diagonal matrix with the vector x on the diagonal, and the diagonal of C_p is S_p^2

REGRESSION FIT OF COMBINED DATA

Taking the updates of the previous model and combining the calculated data with the measured data, additional regression fits were made. Table 9 summarizes the data available for the eight regression fits. Eleven calculated data points were produced for four of the eight categories and seven points for the Hanford dataset. No reactor physics calculations were completed for the VVER-440, VVER-1000, or the RBMK-1000 reactor types, however the dataset required at least one low burnup point in each category, so one

data point was added to obtain reasonable results at low burnup. The lowest burnup point from all the calculated data was used in all eight categories.

Reactor Type (Moderator)	Measured Data	Calculated Data
CANDU (Heavy Water)	3	11
GE (BWR)	78	11
Hanford (Graphite)	46	7
Magnox (Graphite)	4	11
RBMK (Graphite)	41	1
Westinghouse (PWR)	156	11
VVER-1000 (PWR)	14	1
VVER-440 (PWR)	41	1

Table 9: Complete Datasets from Measurements and Calculated Data

Figure 30 to Figure 37 display the eight regression fits with the corresponding data. In all cases, the high burnup points are from fuel rod measurements, while the low burnup points are from physics-model calculations. The red lines display 2-sigma error bounds of proportional error. Note that these error bounds do not include regression-fit or detection level uncertainty.

Figure 38 compares the eight regression fits to each other and provides a visual measure as to how well the classification algorithm should work; At some burnups, the regression lines are well separated from each other which is an indication that classification will work well at this burnup. At other burnups, specifically low burnup, the regression lines are not well separated, and classification may be less certain at these levels, resulting in a potential for higher miscalculation probabilities. Also present on these plots are error

bars that represent error in measured isotopic ratios (proportional and detection-level).

Note that measurement error gets large at small burnups.

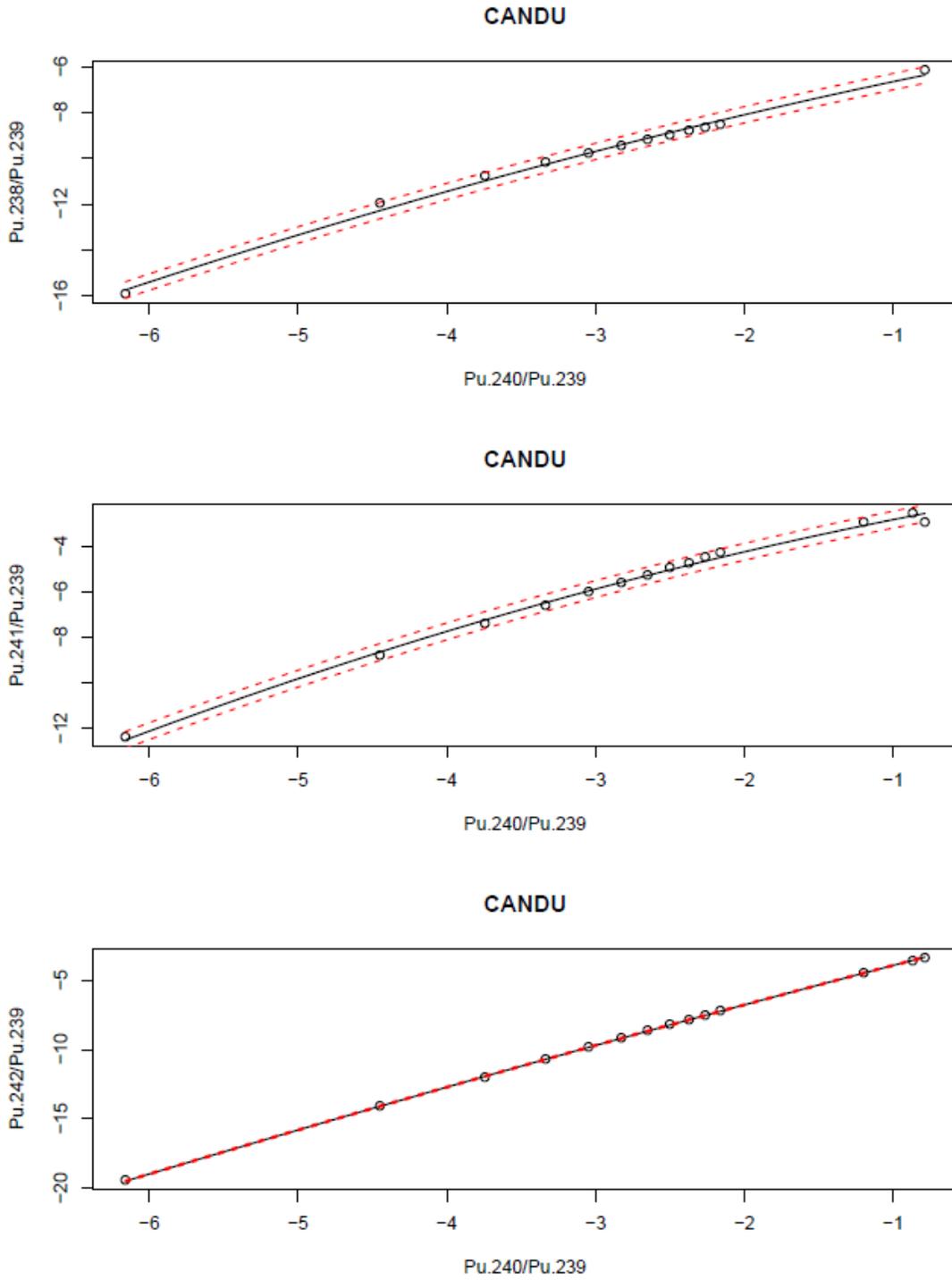


Figure 30: Regression Fit and 2-sigma error bounds of CANDU (Heavy Water) Reactor Type with Measured and Calculated Data

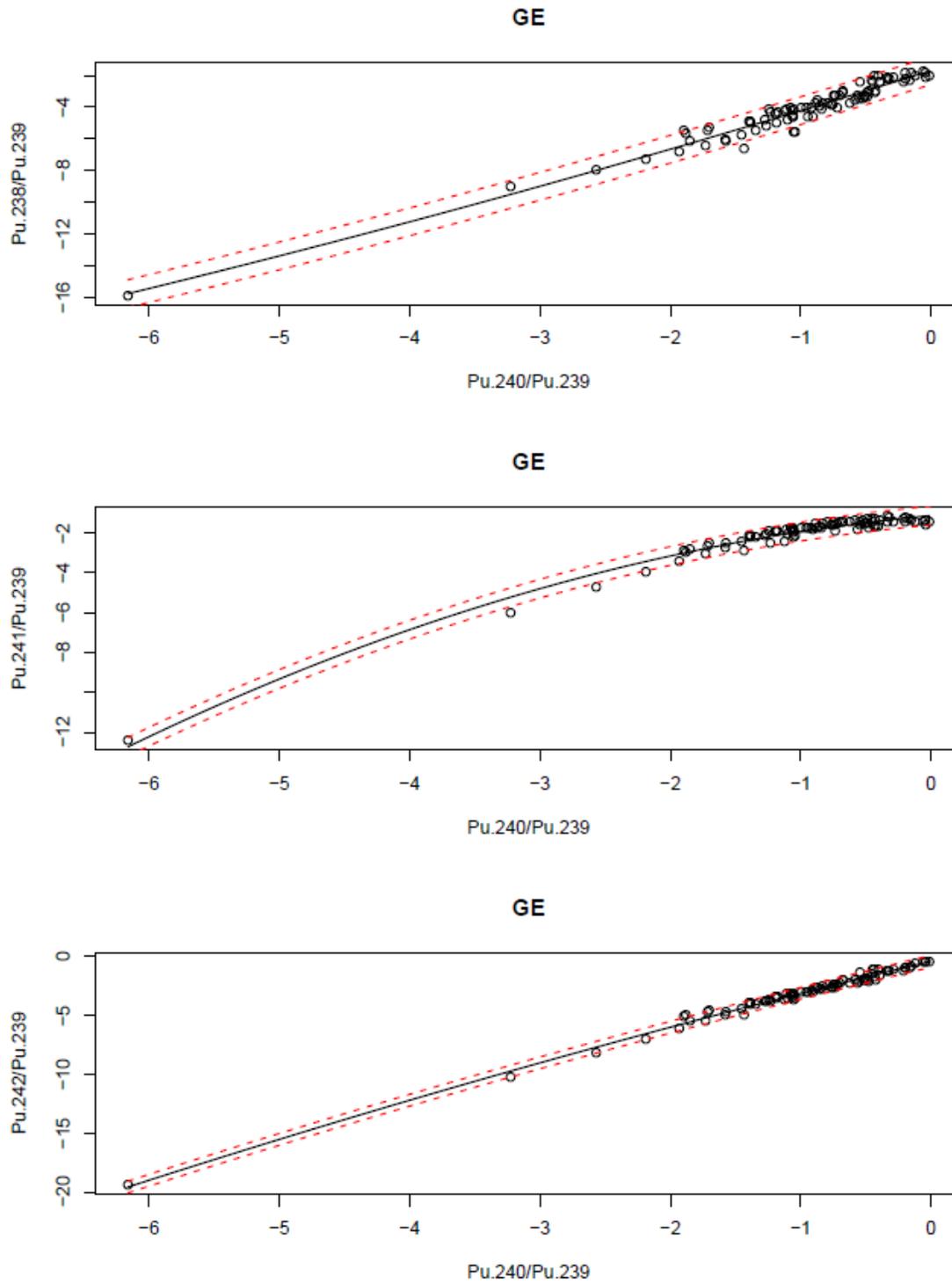


Figure 31: Regression Fit and 2-sigma error bounds of GE (BWR) Reactor Type with Measured and Calculated Data

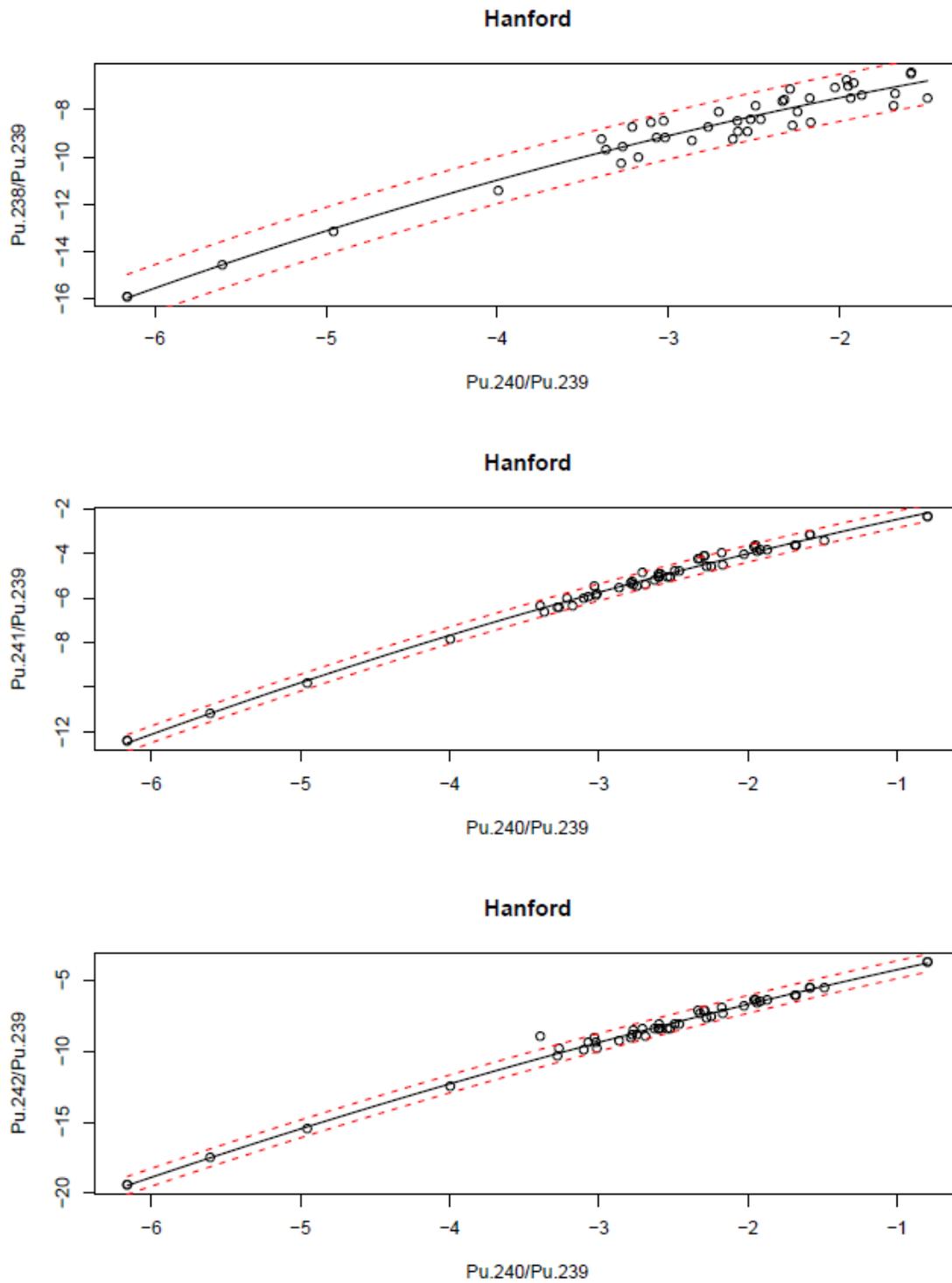


Figure 32: Regression Fit and 2-sigma error bounds of Hanford (Graphite) Reactor Type with Measured and Calculated Data

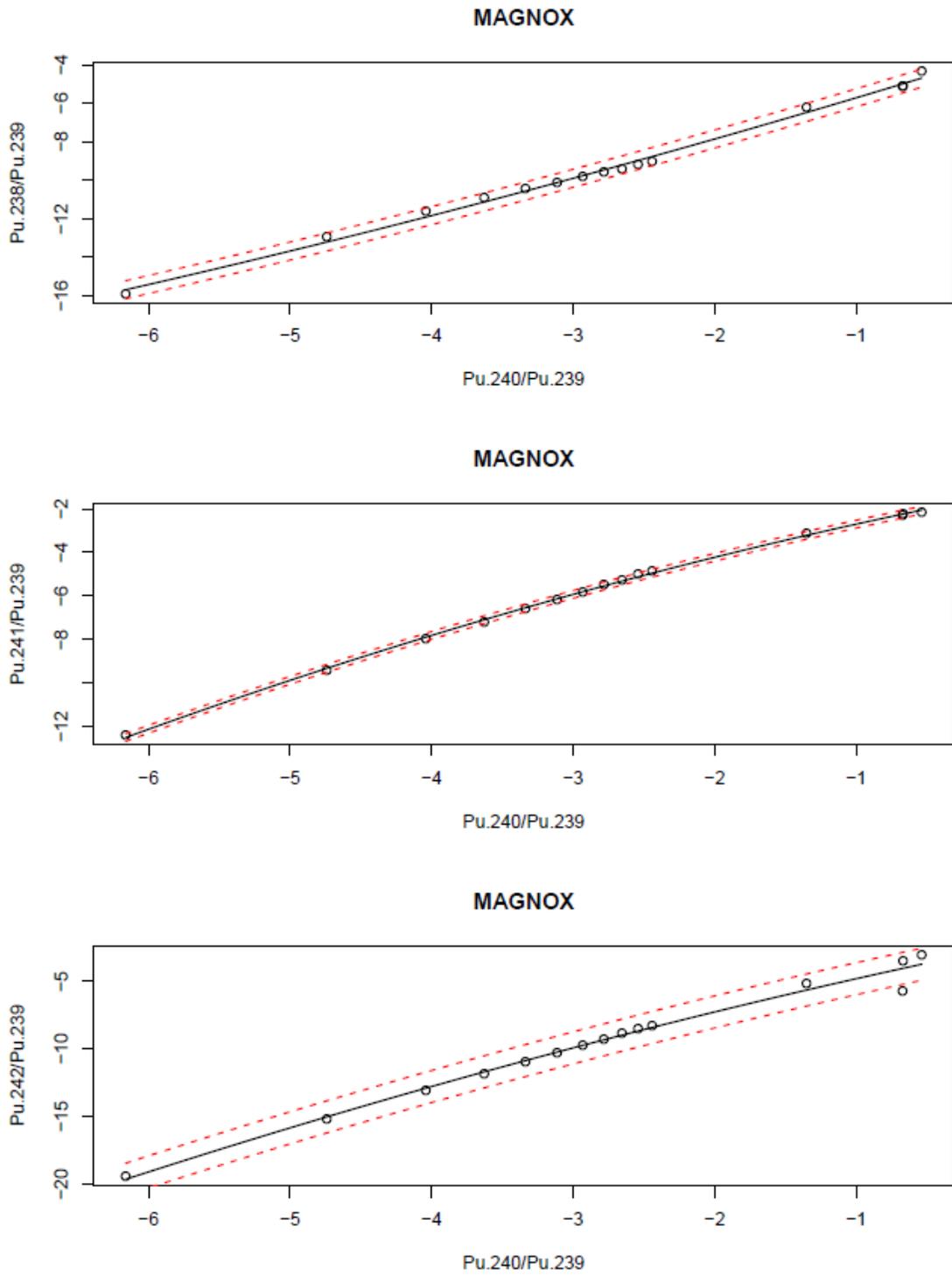


Figure 33: Regression Fit and 2-sigma error bounds of Magnox (Graphite) Reactor Type with Measured and Calculated Data

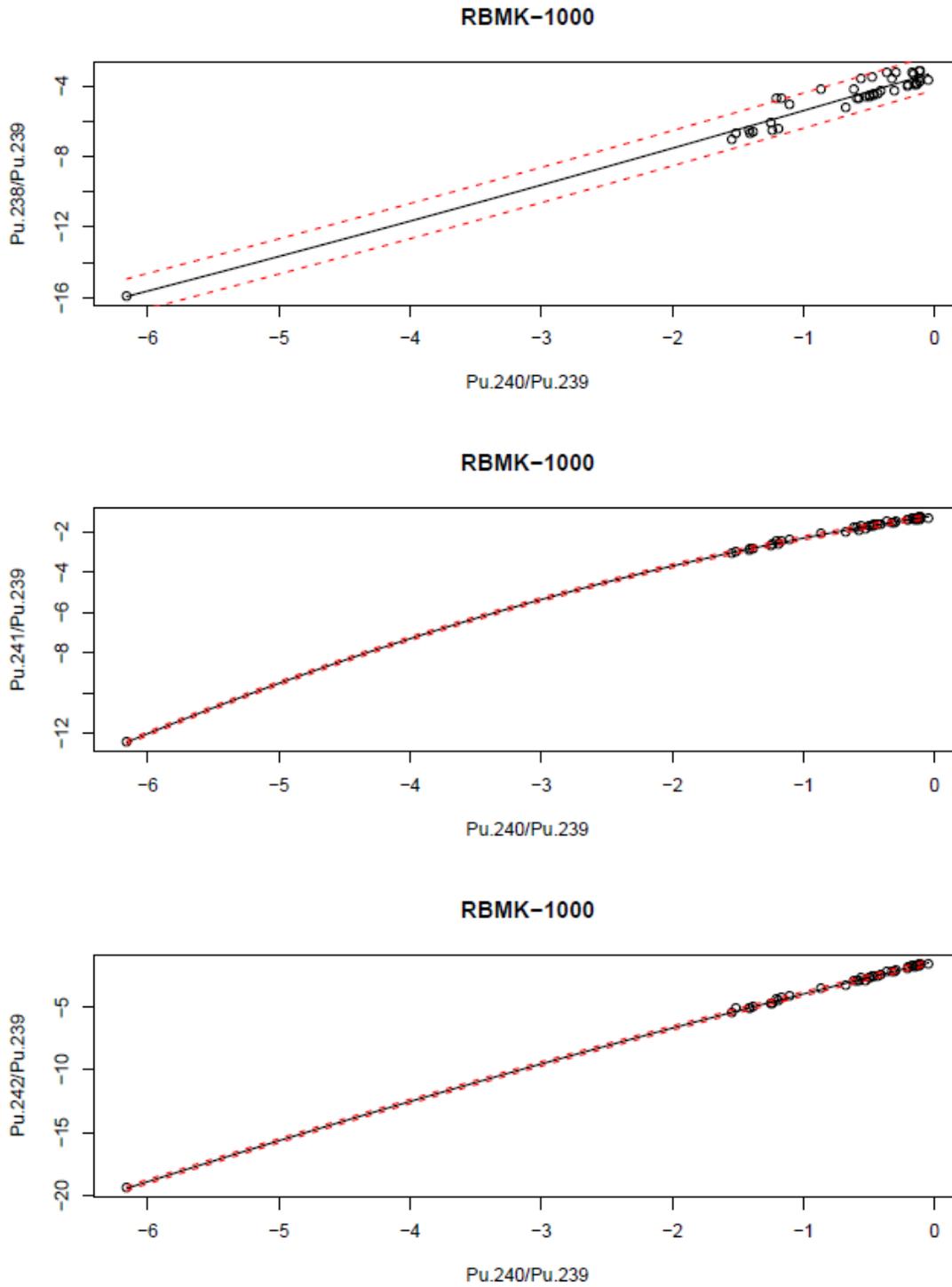


Figure 34: Regression Fit and 2-sigma error bounds of RBMK-1000 (Graphite) Reactor Type with Measured and Calculated Data

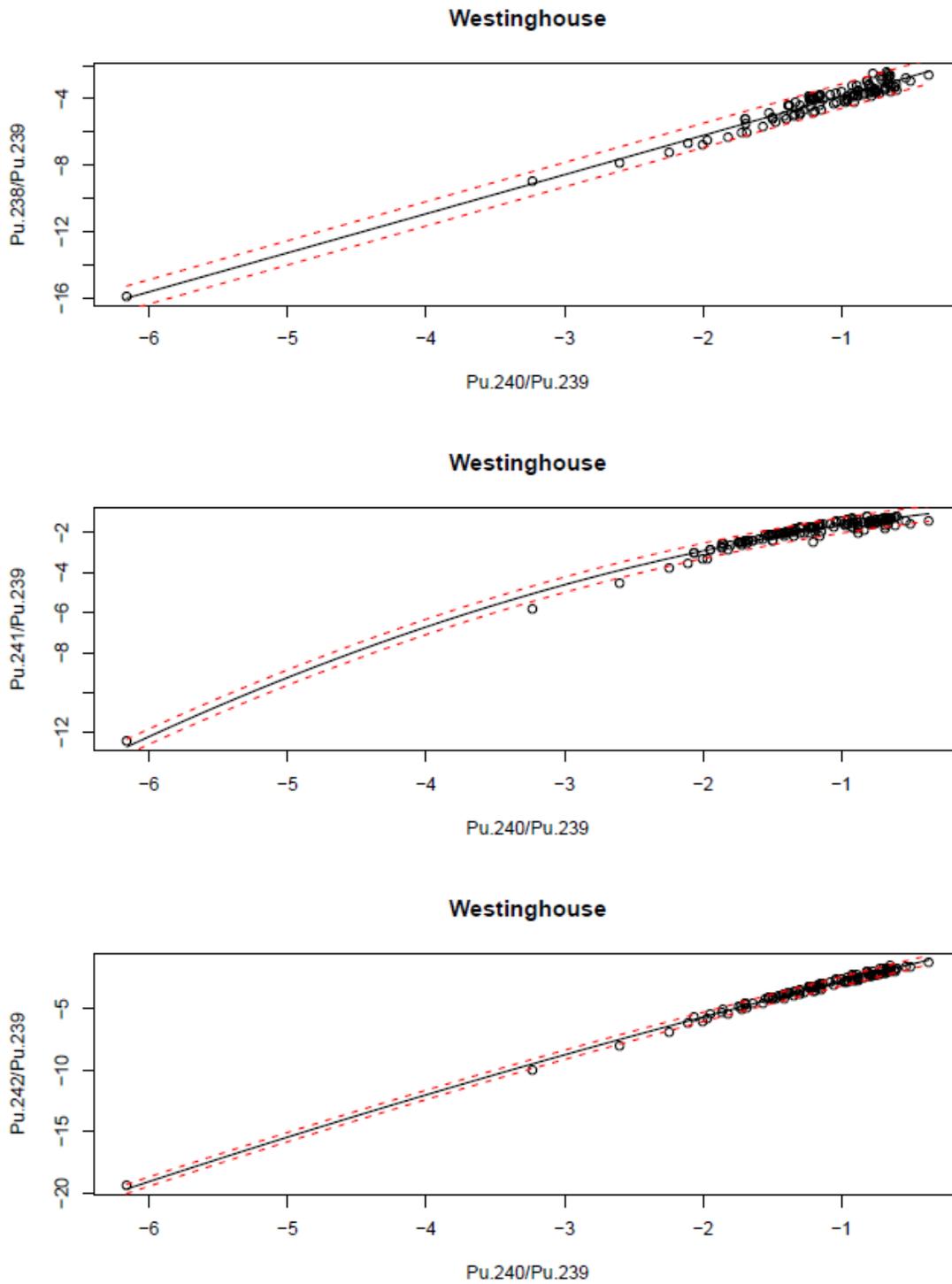


Figure 35: Regression Fit and 2-sigma error bounds of Westinghouse (PWR) Reactor Type with Measured and Calculated Data

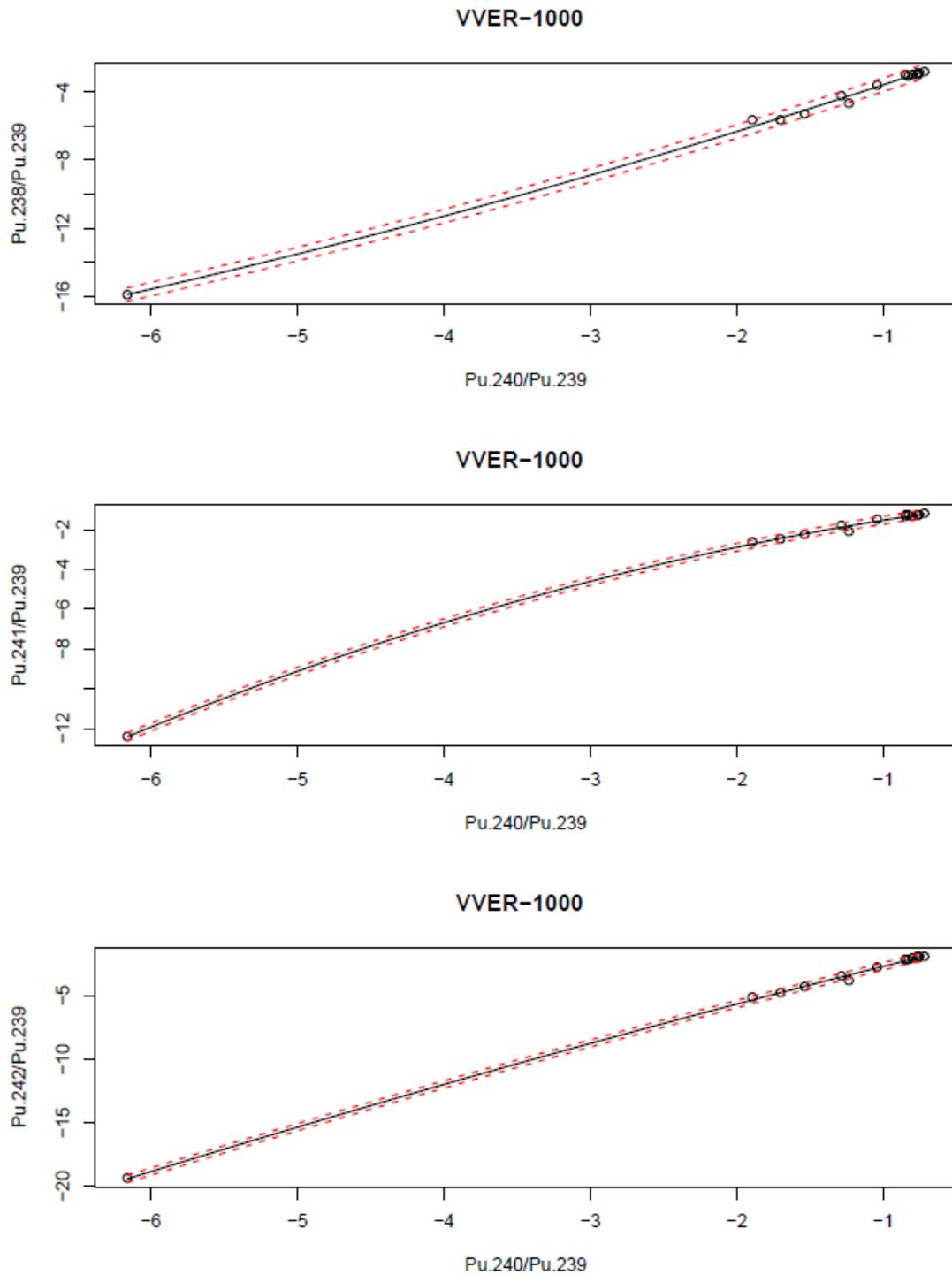


Figure 36: Regression Fit and 2-sigma error bounds of VVER-1000 (PWR) Reactor Type with Measured and Calculated Data

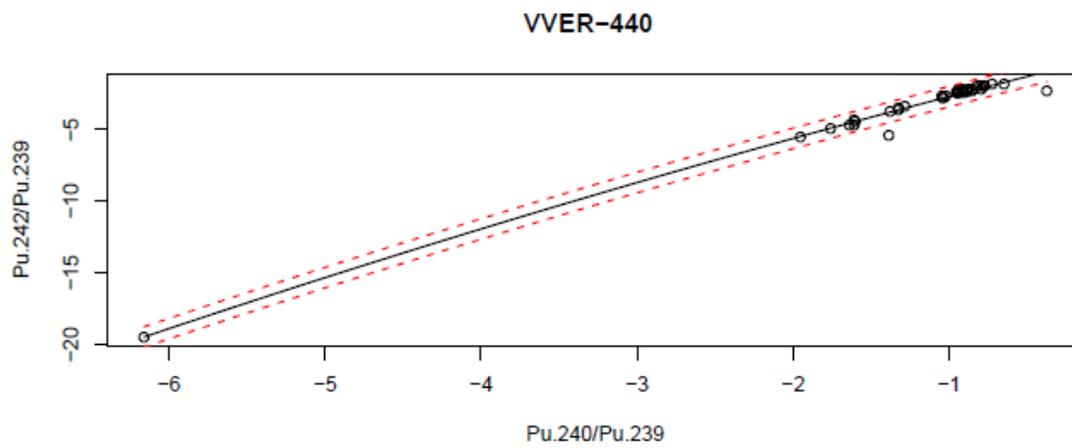
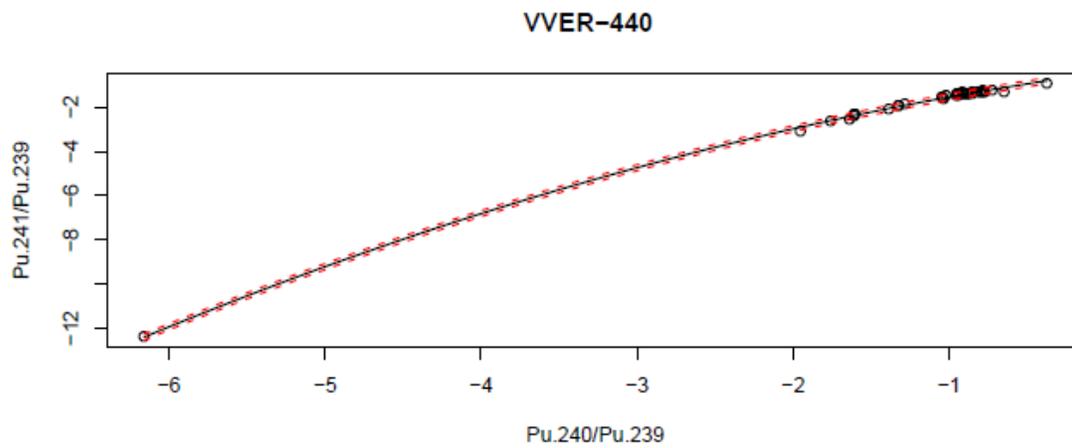
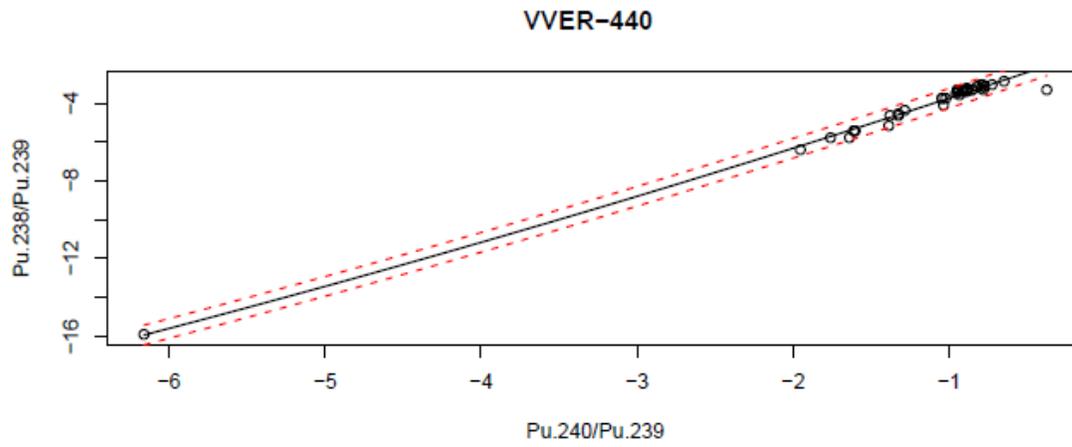


Figure 37: Regression Fit and 2-sigma error bounds of VVER-440 (PWR) Reactor Type with Measured and Calculated Data

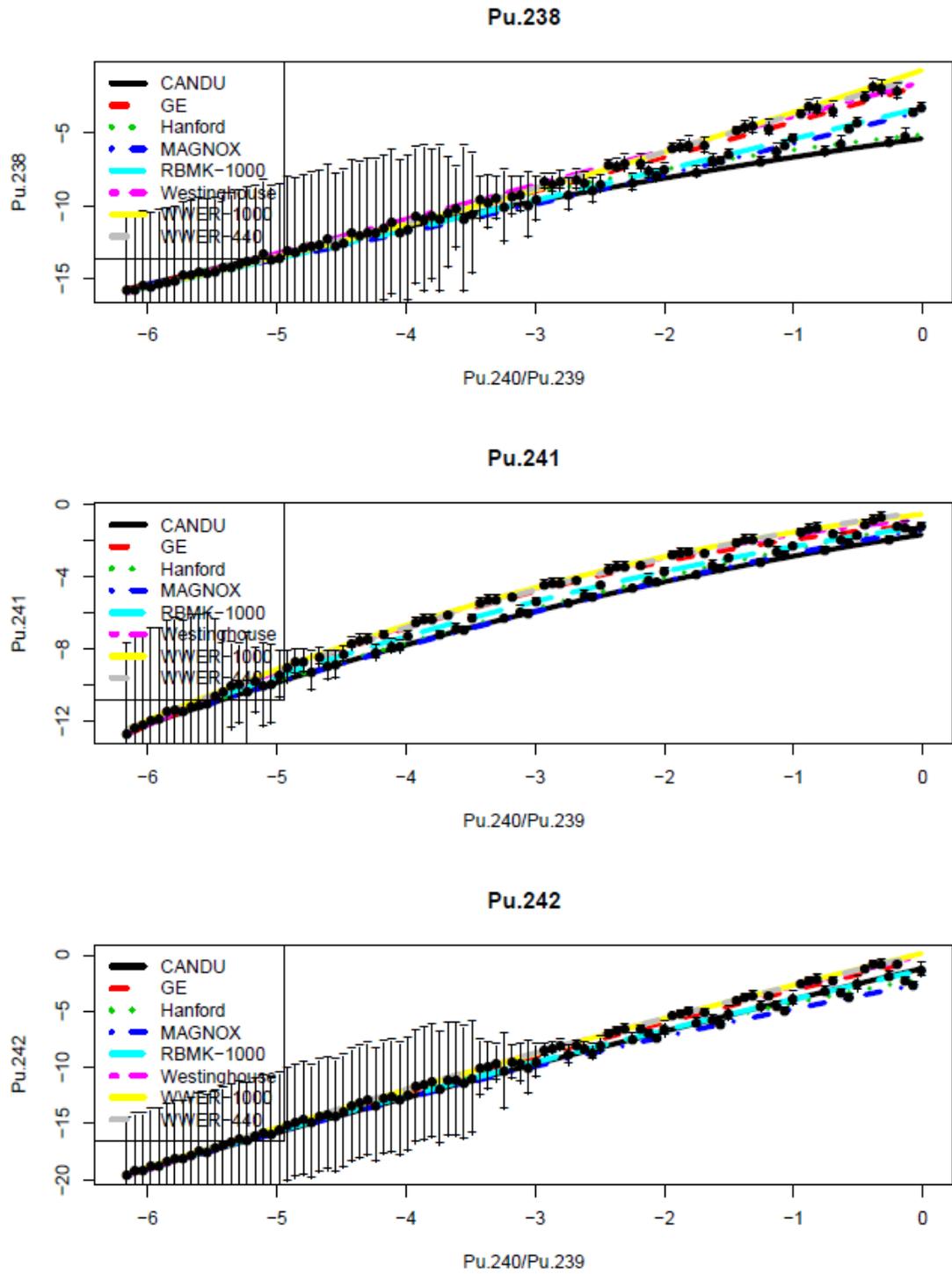


Figure 38: Comparison of Regression Fits with Classification Error for the Eight Reactor Types

With the regressions completed including the expanded dataset incorporating the measured and calculated data, the performance of the classification algorithm can be tested and compared to the previous version. Table 10 shows the Average Posterior matrix like what was calculated in Table 8. Note that there are now eight categories since two new reactor types were added. The results in Table 10 have been calculated using only measured data and all of it is used for training. Note that the results are not dramatically different than seen in the previous version.

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	Westinghouse	VVER-1000	VVER-440
CANDU	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hanford	0.108	0.998	0.062	0.118	0.021	0.000	0.000	0.000
MAGNOX	0.000	0.000	0.804	0.000	0.000	0.000	0.000	0.000
RBMK-1000	0.000	0.001	0.008	0.954	0.002	0.000	0.000	0.000
GE	0.000	0.026	0.000	0.007	0.734	0.109	0.093	0.033
Westinghouse	0.000	0.197	0.000	0.000	0.150	0.508	0.199	0.096
VVER-1000	0.000	0.062	0.000	0.000	0.065	0.092	0.776	0.071
VVER-440	0.000	0.032	0.000	0.000	0.032	0.184	0.358	0.407

Table 10: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types

Given the wider range of burnups in the Version 3 regression fits and the nature of the model variability, it makes sense to calculate the Posterior Classification Matrix as a function of burnup. Table 11, Table 12, and Table 13 calculate the Posterior Matrix for low (.01), median (0.3) and high (1) burnup, with burnup represented by $^{240}\text{Pu}/^{239}\text{Pu}$. The low and high values represent approximately the low and high values in the dataset, with "median" representing the dataset median. Note that burnup strongly affects classification errors, with classification errors highest at low burnup and the errors improving for higher burnup. It should also be noted that the posterior matrix is not calculated using a validation

dataset, but by employing a theoretical integration using the multi-variate normal model that is the basis for the classification algorithm. Such a calculation may produce optimistic results because it does not account for any inadequacies in the model.

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	Westinghouse	VVER-1000	VVER-440
CANDU	0.465	0.098	0.229	0.079	0.055	0.025	0.013	0.036
Hanford	0.097	0.366	0.123	0.165	0.091	0.059	0.025	0.074
MAGNOX	0.223	0.111	0.378	0.074	0.082	0.049	0.028	0.055
RBMK-1000	0.079	0.156	0.081	0.269	0.127	0.107	0.055	0.126
GE	0.052	0.091	0.074	0.121	0.209	0.175	0.117	0.162
Westinghouse	0.026	0.057	0.043	0.106	0.174	0.245	0.188	0.161
VVER-1000	0.016	0.028	0.027	0.054	0.108	0.185	0.440	0.142
VVER-440	0.040	0.080	0.055	0.135	0.159	0.154	0.141	0.236

Table 11: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.01$

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	Westinghouse	VVER-1000	VVER-440
CANDU	0.994	0.006	0.000	0.000	0.000	0.000	0.000	0.000
Hanford	0.003	0.815	0.064	0.102	0.018	0.000	0.000	0.000
MAGNOX	0.000	0.073	0.927	0.000	0.000	0.000	0.000	0.000
RBMK-1000	0.000	0.086	0.000	0.911	0.003	0.000	0.000	0.000
GE	0.000	0.018	0.000	0.002	0.702	0.201	0.024	0.052
Westinghouse	0.000	0.000	0.000	0.000	0.189	0.558	0.125	0.128
VVER-1000	0.000	0.000	0.000	0.000	0.021	0.123	0.744	0.112
VVER-440	0.000	0.000	0.000	0.000	0.057	0.116	0.110	0.717

Table 12: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.3$

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	Westinghouse	VVER-1000	VVER-440
CANDU	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Hanford	0.000	0.998	0.002	0.000	0.000	0.000	0.000	0.000
MAGNOX	0.000	0.001	0.982	0.017	0.000	0.000	0.000	0.000
RBMK-1000	0.000	0.000	0.017	0.982	0.000	0.000	0.000	0.000
GE	0.000	0.000	0.000	0.001	0.810	0.177	0.008	0.004
Westinghouse	0.000	0.000	0.000	0.000	0.187	0.767	0.012	0.034
VVER-1000	0.000	0.000	0.000	0.000	0.004	0.009	0.984	0.003
VVER-440	0.000	0.000	0.000	0.000	0.005	0.034	0.004	0.957

Table 13: Average Posteriors Developed from SFCOMPO/NF Data for all Eight Reactor Types for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 1$

Figure 39 plots misclassification error for each reactor series as a function of burnup. Note the error becomes large as burnup goes to zero but improves as burnup increases. The vertical line on the plot identifies median burnup. It is also important to note that the classes RBMK-1000, MAGNOX, Hanford, and CANDU seem to have the lowest classification error, which is expected since the moderator types and operating parameters of these reactors is sufficiently different.

DISCUSSION AND NEXT STEPS

Based on the results of the from this interim version of the classification algorithm, the following conclusions and improvements were identified for the next version of the algorithm:

- Calculated Data – The calculated data fits relatively well with the measured data since the ORIGEN-ARP models are validated within the expected operating conditions for power reactors. As previously noted, it will be important to determine the uncertainty in the reactor physics models and how it transfers to the overall uncertainty of the calculated data so that it can be properly weighted in the classification algorithm. Adding uncertainty will likely make the data more

disperse, creating more overlap in the data and less precision in the probability results of the tool. For this interim version the quantity of calculated data was a small percentage of the total samples and as noted in the Regression Model Updates section, the uncertainty of the calculated data was assumed to be zero for all data points.

- Data in General – Additional measured data for the CANDU, MAGNOX, RBMK and VVER reactors that expand on the expected operational parameters and representing a wider range of conditions would make the classification algorithm more robust. Additionally, more data for all reactor types in the low burnup region would help lower the classification error shown in Figure 39.
- Consolidation of similar reactor types – Ultimately it may be important to be able to identify a specific reactor type, or even specific reactor, from plutonium ratios, but with insufficient data, it may be beneficial to combine similar reactor types into one category. Since there is not a qualified data source available to expand the dataset, it makes sense to combine the Westinghouse, VVER-440 and VVER-1000 into a “PWR” reactor type.
- Outlier Category – It is important to include an outlier category, so that the classifier does not error out due to numerical underflow and a future user knows when the classification model does not fit the data. As we continue to develop the classification tool, this will be an important addition.

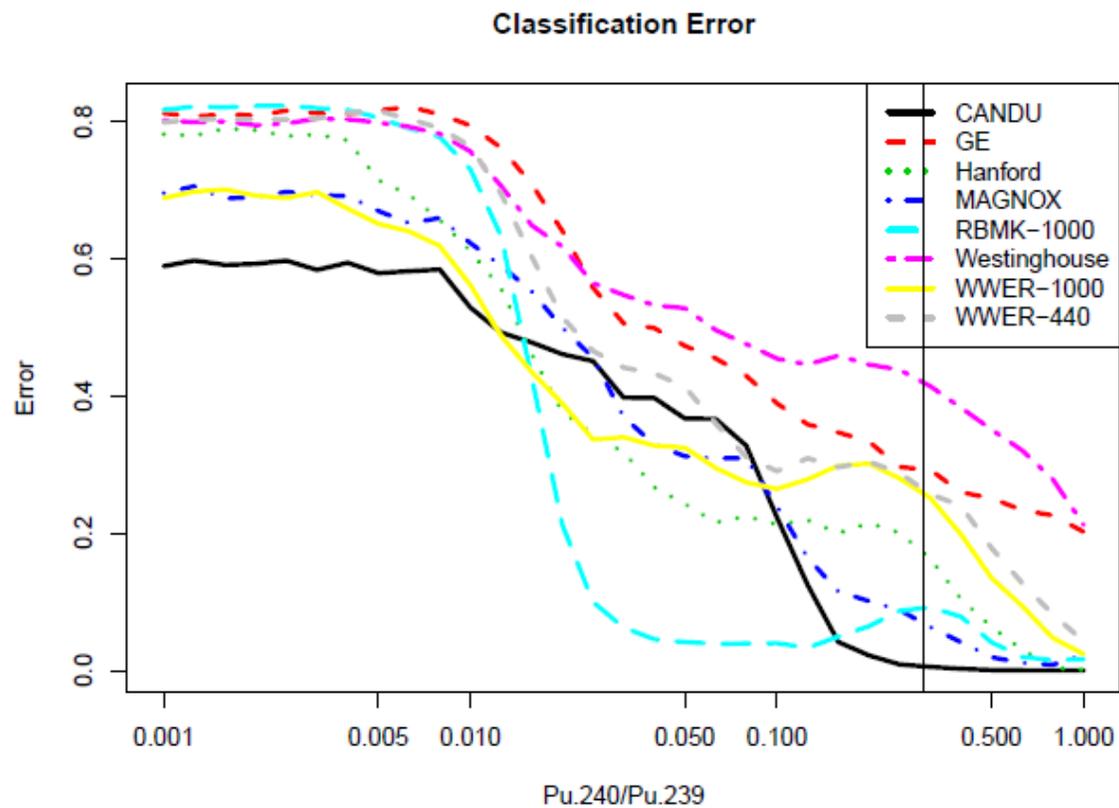


Figure 39: Error rates as a function of Burnup ($^{240}\text{Pu}/^{239}\text{Pu}$) for the different reactor types

Multivariate Characterization with Consolidated Reactor Types (Version 2)

Implementing some of the identified improvements from the previous classification algorithm, it was determined that consolidating the reactor types with similar operating parameters could help the performance. This version builds on the previous interim version with the incorporation of calculated data with the measured data and combines the Westinghouse, VVER-440 and VVER-1000 categories into a single “PWR” category. As previously stated, all datapoints assumed an uncertainty value of 0% for the generation of the covariance matrices. Figure 40 shows the updated regression fit plots for the consolidated PWR category. The other plots have remained the same (Figure 30 to Figure 34) and can be found in the previous section. Figure 41 shows the updated comparison of regression fits for the six categories.

The values in Table 5 and Table 6 were updated with the consolidated reactor types and the included calculated data. Table 14 and Table 15 show the updated Alpha Parameter and Covariance Matrix Estimates from Regression Fit. These tables were used to develop an excel workbook to run the classification algorithm against a set of plutonium ratios. These results will be shown in a later section.

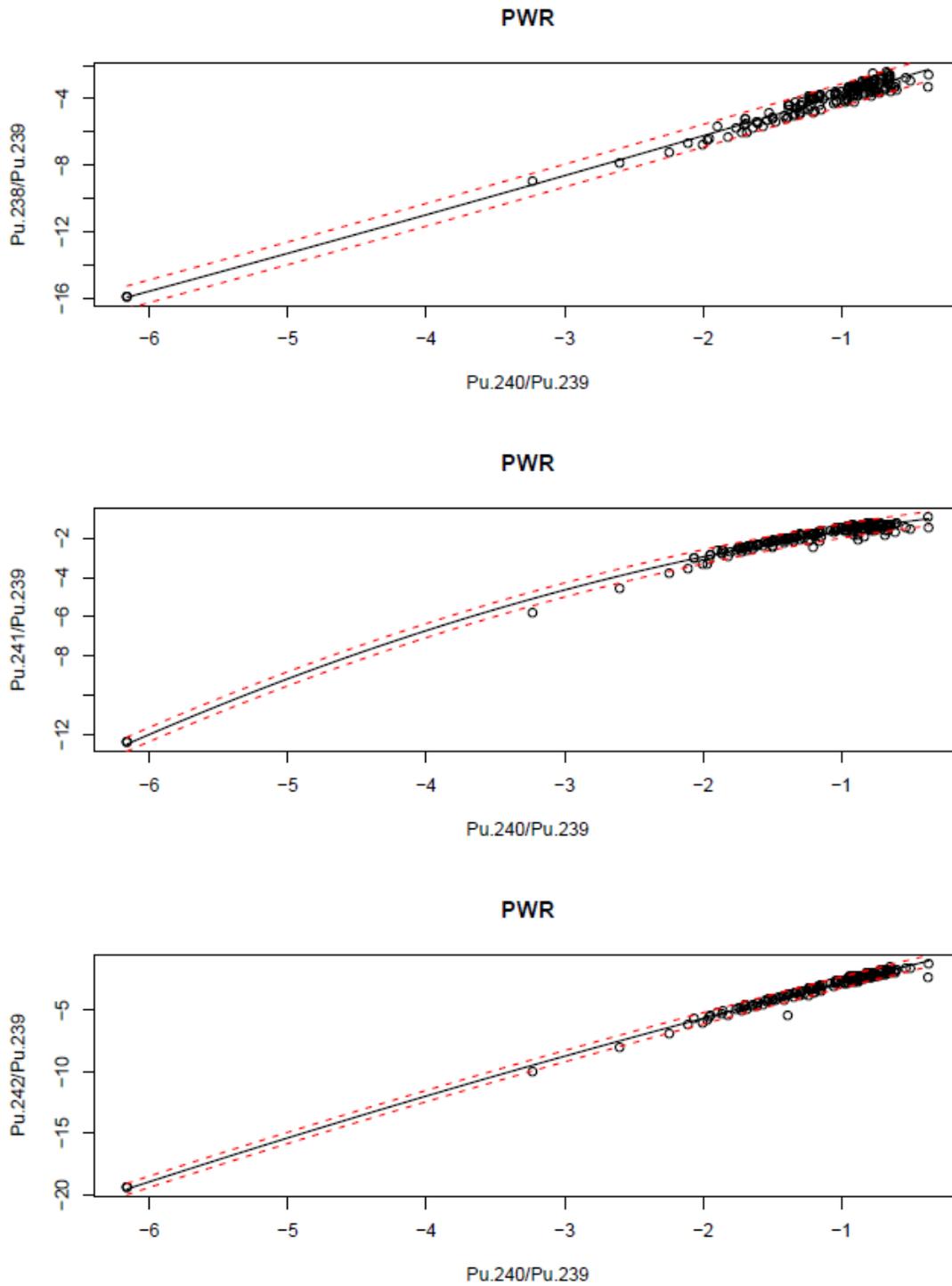


Figure 40: Regression Fit and 2-sigma error bounds of the Consolidate PWR Reactor Type with Measured and Calculated Data

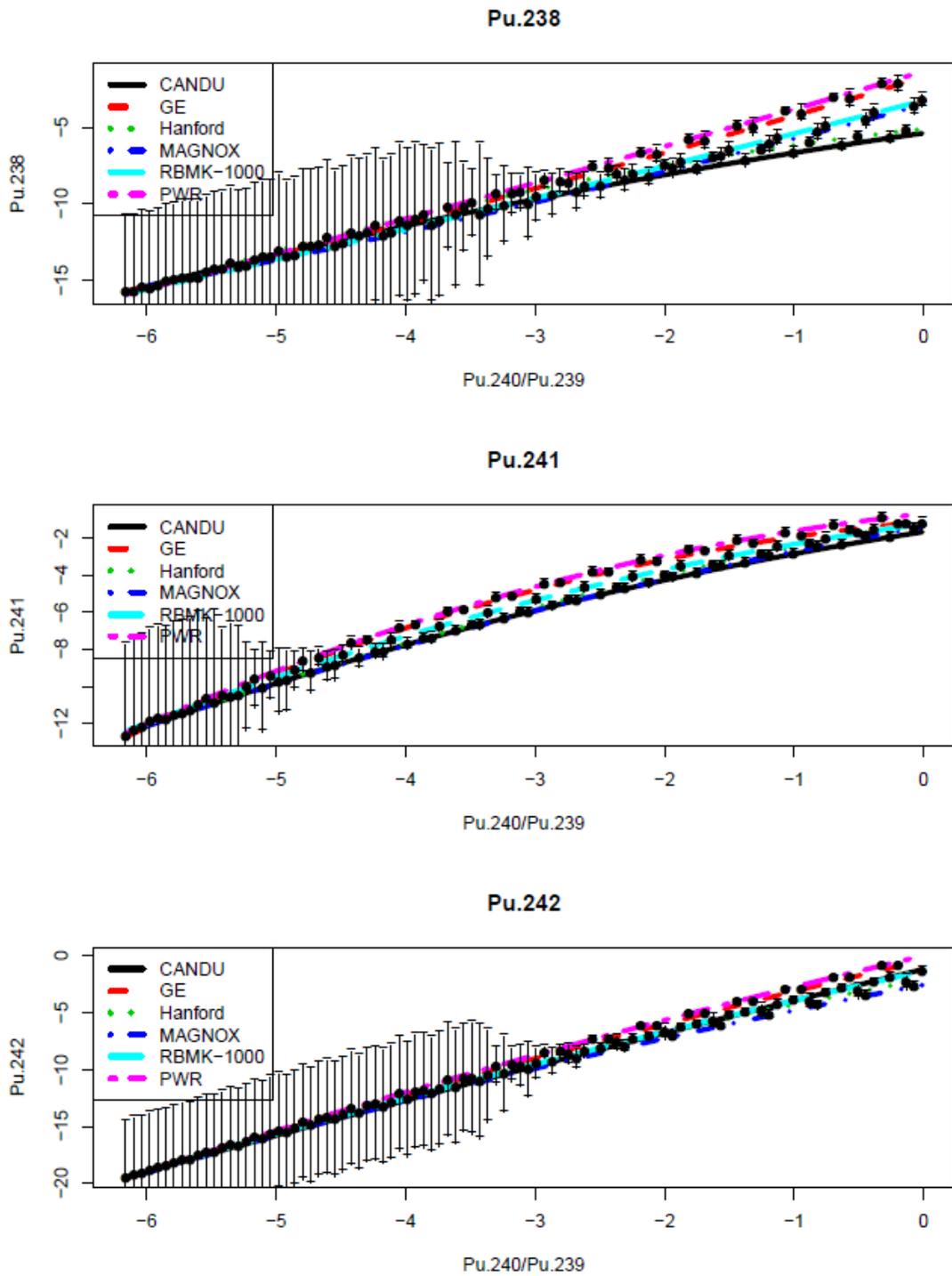


Figure 41: Comparison of Regression Fits with Classification Error for the Six Reactor Types

j Isotopic Ratio	α_{1ij}	α_{2ij}	α_{3ij}
CANDU (i=1)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-5.36387	1.218365	-0.07578
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.65312	1.082261	-0.11127
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-1.17076	2.719197	-0.04112
BWR (i=2)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.69421	2.5760113	0.047003
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.11996	0.5956616	-0.2092
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-0.38217	2.6580769	-0.07435
Hanford (i=3)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-5.09915	0.92552	-0.13603
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.10664	1.251023	-0.09743
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-2.08476	2.071997	-0.12125
MAGNOX (i=4)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-3.47827	2.291904	0.050139
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.3521	1.254888	-0.09043
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-2.53673	2.157641	-0.09995
RBMK-1000 (i=5)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-3.22157	2.1961705	0.021492
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.21277	0.9672895	-0.13835
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-1.33958	2.5251331	-0.06752
PWR (i=6)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.32174	2.4967816	0.019919
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-0.67992	0.7397514	-0.19178
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	0.02367	2.6739753	-0.08158

Table 14: Alpha Parameter Estimates from Regression Fit

	$^{238}\text{Pu}/^{239}\text{Pu}$	$^{241}\text{Pu}/^{239}\text{Pu}$	$^{242}\text{Pu}/^{239}\text{Pu}$
CANDU (i=1)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.031827403	-0.029813002	-0.004333687
$^{241}\text{Pu}/^{239}\text{Pu}$	-0.029813002	0.034547026	0.004684451
$^{242}\text{Pu}/^{239}\text{Pu}$	-0.004333687	0.004684451	0.001203521
BWR (i=2)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.19245095	0.01740169	0.07723688
$^{241}\text{Pu}/^{239}\text{Pu}$	0.01740169	0.05526846	0.03107564
$^{242}\text{Pu}/^{239}\text{Pu}$	0.07723688	0.03107564	0.06284208
Hanford (i=3)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.24823887	0.07410551	0.08122217
$^{241}\text{Pu}/^{239}\text{Pu}$	0.07410551	0.03563191	0.03119836
$^{242}\text{Pu}/^{239}\text{Pu}$	0.08122217	0.03119836	0.098336
MAGNOX (i=4)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.05511872	-0.015894612	0.02709316
$^{241}\text{Pu}/^{239}\text{Pu}$	-0.01589461	0.008077273	0.01418814
$^{242}\text{Pu}/^{239}\text{Pu}$	0.02709316	0.01418814	0.35119417
RBMK-1000 (i=5)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.24946013	0.019223538	0.027386961
$^{241}\text{Pu}/^{239}\text{Pu}$	0.01922354	0.002439338	0.003286498
$^{242}\text{Pu}/^{239}\text{Pu}$	0.02738696	0.003286498	0.007913504
PWR (i=6)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.11676202	0.03306564	0.05713494
$^{241}\text{Pu}/^{239}\text{Pu}$	0.03306564	0.03263444	0.02594396
$^{242}\text{Pu}/^{239}\text{Pu}$	0.05713494	0.02594396	0.05295754

Table 15: Covariance Matrix (C_i) Estimates from Regression Fit

To compare with the previous versions of the classification algorithm, the average posteriors are calculated and show in the following tables. Table 16 shows the overall average posteriors for the six different reactor types. Note that there is an approximately 22% error in classification for the new PWR type, with the majority being misclassified as a GE (BWR). This is an expected outcome since both types are light water moderated

reactors. It is likely that some level of misclassification will occur since there is sufficient overlap in the data that can be seen in Figure 15 to Figure 17.

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	PWR
CANDU	1.000	0.000	0.000	0.000	0.000	0.000
Hanford	0.108	0.691	0.062	0.118	0.021	0.000
MAGNOX	0.000	0.196	0.804	0.000	0.000	0.000
RBMK-1000	0.000	0.037	0.008	0.954	0.001	0.000
GE	0.000	0.025	0.000	0.007	0.681	0.287
PWR	0.000	0.012	0.000	0.000	0.213	0.776

Table 16: Average Posteriors Developed for Six Different Reactor Types with a Consolidated PWR type

Table 17 to Table 19 show the change in average posteriors at the different burnup values of 0.01, 0.3, and 1. With the consolidated PWR category, these show better classification probabilities when compared to the previous version in Table 11 to Table 13.

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	PWR
CANDU	0.466	0.096	0.232	0.087	0.068	0.051
Hanford	0.099	0.389	0.131	0.176	0.113	0.093
MAGNOX	0.237	0.121	0.380	0.087	0.100	0.074
RBMK-1000	0.091	0.184	0.097	0.316	0.159	0.153
GE	0.061	0.109	0.097	0.162	0.295	0.275
PWR	0.044	0.098	0.072	0.162	0.273	0.351

Table 17: Average Posteriors Developed for Six Reactor Types with a Consolidated PWR type for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.01$

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	PWR
CANDU	0.994	0.006	0.000	0.000	0.000	0.000
Hanford	0.011	0.812	0.065	0.091	0.020	0.000
MAGNOX	0.000	0.062	0.938	0.000	0.000	0.000
RBMK-1000	0.000	0.085	0.000	0.912	0.003	0.000
GE	0.000	0.019	0.000	0.001	0.730	0.250
PWR	0.000	0.000	0.000	0.000	0.256	0.744

Table 18: Average Posteriors Developed for Six Reactor Types with a Consolidated PWR type for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 0.3$

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	PWR
CANDU	1.000	0.000	0.000	0.000	0.000	0.000
Hanford	0.000	1.000	0.000	0.000	0.000	0.000
MAGNOX	0.000	0.002	0.982	0.017	0.000	0.000
RBMK-1000	0.000	0.000	0.014	0.985	0.000	0.000
GE	0.000	0.000	0.001	0.000	0.801	0.198
PWR	0.000	0.000	0.000	0.000	0.183	0.817

Table 19: Average Posteriors Developed for Six Reactor Types with a Consolidated PWR type for Burnup $^{240}\text{Pu}/^{239}\text{Pu} = 1$

Table 19 shows that discrimination is improved at high burn-up. However, GE and PWR are still confused at a rate of about 20%. Figure 42 shows the updated classification error rates as a function of Burnup ($^{240}\text{Pu}/^{239}\text{Pu}$) for the different reactor types.

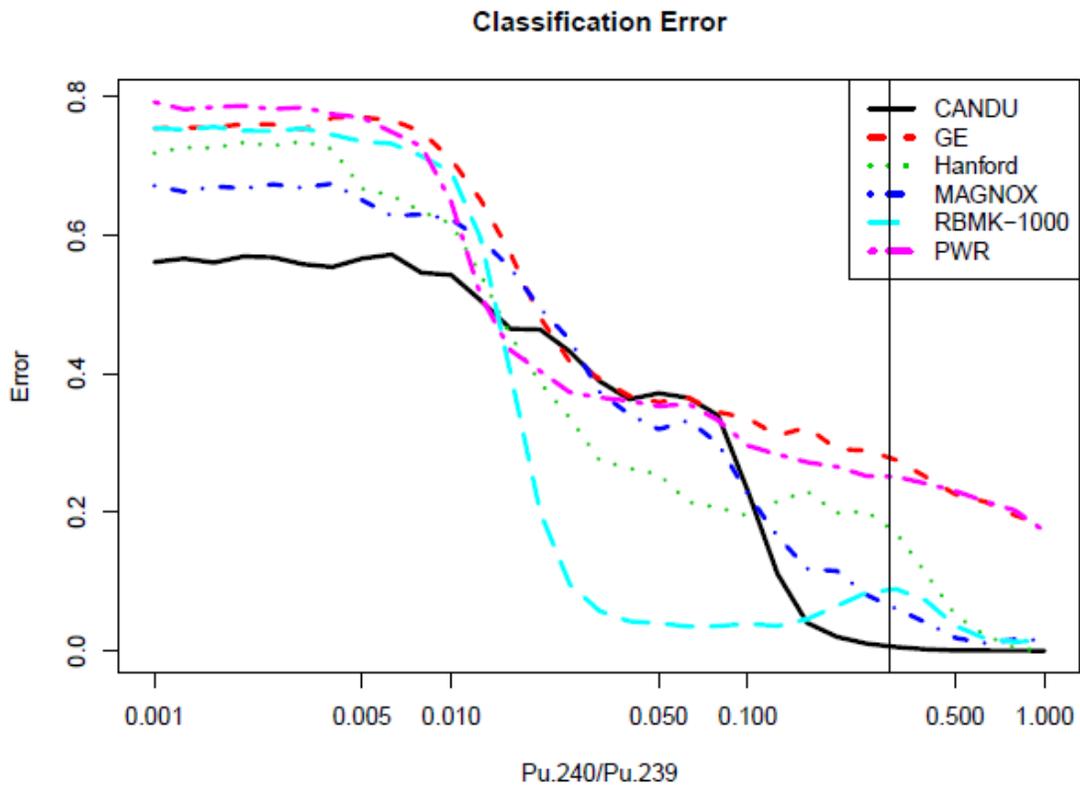


Figure 42: Error rates as a function of Burnup ($^{240}\text{Pu}/^{239}\text{Pu}$) for the different reactor types

Note that the classification error is still highest in the low burnup region where there is significant data overlap, a limited amount of data points, and relatively small differences in isotopics.

Multivariate Characterization with Outlier Category (Version 3)

This section describes the final updates and refinements to the classification algorithm. The major difference is the addition of an “Outlier” category to the discrimination model, with everything else is the same as Version 2. This includes consolidating the measured data from the SFCOMPO/NF dataset with calculated data developed for Version 2 with the addition of the outlier category with a default value of 0% uncertainty was assumed for all data points for the generation of the covariance matrices. This category would fit the data points that lie outside of the range parameters of the reactor types categorized in the model. The inclusion of the outlier category also eliminates division by zero when such data points are fed into the algorithm.

Working with the statistician, the final version of the classification incorporated the Outlier category. The regression fits from the previous version are still used, but now the outlier category assumes a 1% chance of any data point being an outlier. This sets the prior probability associated with the outlier category at 1%.

The outlier distribution is defined to be uniform on the logged data and does not have a specific region identified for the distribution. It will extend as far as needed around the multivariate normal distributions that define the other groups. The magnitude of $f_{Outlier}$ is defined to be as large as the most disperse distribution associated with the other categories:

$$\log(f_{Outlier}) = -0.5(\text{Min}_i \log(|C_i|) + P)$$

Where P represents the dimension of Y' and C_i is the covariance matrix of category i. Since the error E is:

$$E \left((Y' - \mu_i)^T C_i^{-1} (Y' - \mu_i) \right) = P$$

when Y' is from category i , it shows that the above choice for $f_{Outlier}$ will produce an outlier loglikelihood that is roughly the same size as the most disperse category.

Since an additional Outlier category has been added, the prior distribution now contains an Outlier component, and the prior probability associated with Outliers should be set to the frequency that outliers are expected to occur in the data. The default value that was chosen is 1% and this roughly translates into a level of significance of 1%. That is, when Y' is not an Outlier, the posterior probability for the outlier category is 1% on average. As more data is incorporated into the training set, this number may need to be modified in the future.

To generate Table 20, 35 outlier points were produced. Five outliers were produced from draws of a uniform distribution, which consequently, these five data points were easily seen to be outliers. The other 30 outliers were produced by averaging three data points from the same reactor series together, and these points were not readily identified as outliers with only six points classified as such. Note that with this change to the algorithm, approximately 5% of the legitimate points are now being classified as outliers, and the reactor type categories with less data points have an increased probability of a data point being classified as an outlier.

	CANDU	Hanford	MAGNOX	RBMK-1000	GE	PWR	Outlier
CANDU	0.897	0.000	0.000	0.000	0.000	0.000	0.103
Hanford	0.107	0.674	0.061	0.116	0.020	0.000	0.022
MAGNOX	0.000	0.168	0.704	0.000	0.000	0.000	0.127
RBMK-1000	0.000	0.035	0.008	0.948	0.001	0.000	0.008
GE	0.000	0.020	0.000	0.006	0.635	0.280	0.059
PWR	0.000	0.000	0.000	0.000	0.206	0.765	0.029
Outlier	0.000	0.095	0.020	0.028	0.254	0.313	0.290

Table 20: Average Posteriors Developed for Six Different Reactor Types and an Outlier Category

Summary and Development of a Simple Tool for Reactor Type of Origin Classification

With *Version 3* of the classification algorithm complete, an excel workbook was developed to incorporate the parameters generated by the regression algorithms for the three different versions that were developed. All three versions are included for comparison purposes to show how the model developed. Using the alpha parameter and covariance matrix estimates from the regression fits, the posterior probability from the input isotopic ratios can be calculated. The steps in the calculation are summarized below, and an example is provided based on the *Version 1* of the algorithm. The steps are the same for *Versions 2 and 3*. The sample data for this is taken from the Fukushima-Daini-2 plant, which is a GE design. The original ratios are input into Y below. Appendix A shows the tables and calculations completed in the excel workbook in the “behind the scenes” tabs for *Version 3* of the algorithm.

First, transform the original data, represented by Y , into the transformed values Y' and B' .

$$Y = (0.0101279, 0.3887468, 0.17289, 0.0485934)$$

The Transformed Burnup, B' is given by:

$$B' = \log(Y2) = \log\left(\frac{{}^{240}\text{Pu}}{{}^{239}\text{Pu}}\right)$$

and the Transformed Isotopic Ratios are, Y' , are given by:

$$Y' = \log(Y1, Y3, Y4) = \log\left(\frac{{}^{238}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{241}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{242}\text{Pu}}{{}^{239}\text{Pu}}\right)$$

Therefore:

$$Y' = (-4.5924636, -1.7550996, -3.0242686)$$

$$B' = -0.944827$$

Next, calculate the loglikelihood for each category, i :

$$lglik_i = -0.5[\log(|C_i|) + (Y - \mu_i)^T C^{-1}(Y - \mu_i)]$$

With

$$\mu_i = \alpha_{1i} + \alpha_{2i}B'$$

Then calculate the posterior distribution $p(i|Y)$ using the formulas:

$$Q_i = \exp(lglik_i) \times prior_i$$

And finally

$$p(i|Y) = \frac{Q_i}{\sum_{i=1}^6 Q_i}$$

Combining the data from Table 5 and Table 6 with the above equations, the following terms can be calculated, and finally the posterior (final column in Table 21)

	log(C.i)	Quadratic.i	lglik.i	Q.i	p(i Y)
GE	-8.862	2.769	3.047	3.507	0.912
Hanford	-7.977	13.502	-2.763	0.011	0.003
RBMK-1000	-13.450	210.443	-98.497	0.000	0.000
Westinghouse	-11.629	10.278	0.675	0.327	0.085
VVER-1000	-15.030	40.499	-12.734	0.000	0.000
VVER-440	-11.408	27.807	-8.200	0.000	0.000

Table 21: Calculation of Posterior Probability from Isotopic Ratios obtained from Fukushima-Daini-2

This shows a probability of approximately 91% that the plutonium sample came from a GE reactor, and 8% probability it came from a Westinghouse design. The high probability is expected since this measured data point was part of the training data for the

regression. Additionally, since the GE and Westinghouse reactors utilize light water as a moderator, it is reasonable to expect some probability in classifying the data as that reactor type.

Figure 43 shows the main data entry sheet for the Classification Tool excel workbook. On this sheet, isotopic ratios are entered in the green section. Once the data is entered, clicking on the “Calculate Posterior Probability” button will run a macro that calculates the Posterior Probability for the three different version of the classification algorithm using the data other sheets in the workbook. In the figure, isotopic ratios were taken from a BWR dataset, and the macro was run. Note the change in probability from the different versions, based on the changes made.

these assumptions, a Monte Carlo analysis was used to generate a set of 300 possible variations of the measured values for the set of isotope ratios (1200 total generated isotope ratios). Then each set was run through the classification tool to generate results for each of these inputs. The 300 results from the tool are averaged to provide a single probability result, and the variance in the data is also reported. Research by ORNL for the NRC showed that measured isotopic data has a normal distribution, especially when completed by the same organization. (Gauld I., 2011) Therefore, this informal method to generate a normal distribution on the input sample data allows the user to identify large variability in the classification results and provide confidence in the reactor type.

To demonstrate how this works within the tool, a datapoint that represents a PWR reactor type was selected and three different uncertainties in the datapoint were chosen (1%, 5%, and 10%) to determine if the approach was reasonable. In a real-world scenario, only one sample set with its corresponding uncertainty value would be used in the tool. Figure 44 to Figure 47 show sample distributions that were generated for the four different isotope ratios with 300 data points that were then analyzed by the classification tool.

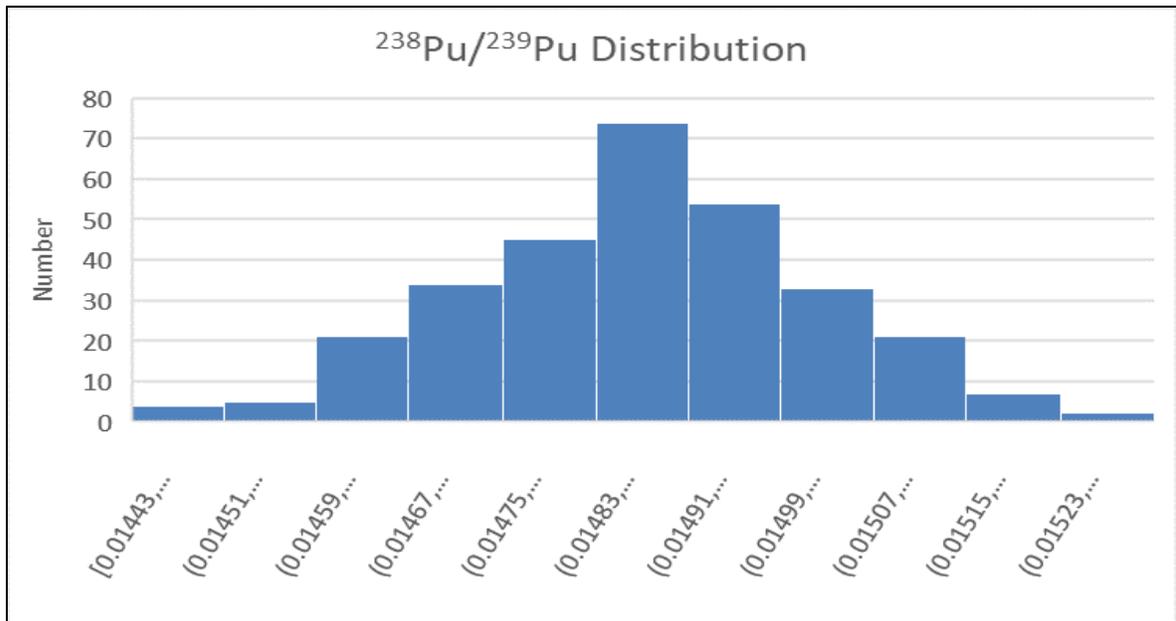


Figure 44: Distribution of values generated for the $^{238}\text{Pu}/^{239}\text{Pu}$ ratio

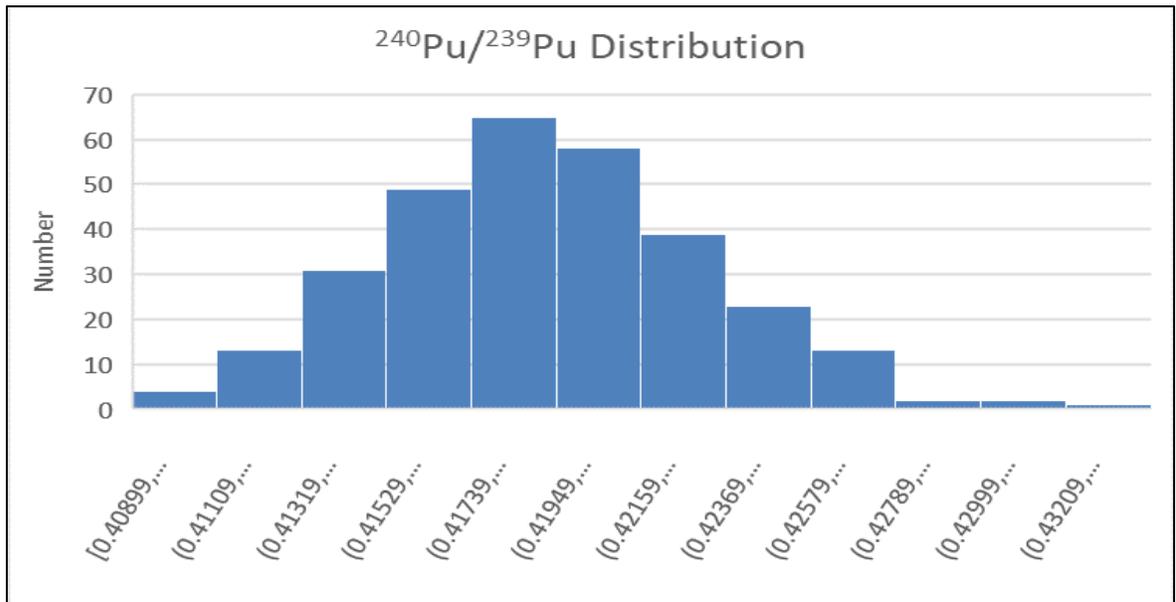


Figure 45: Distribution of values generated for the $^{240}\text{Pu}/^{239}\text{Pu}$ ratio

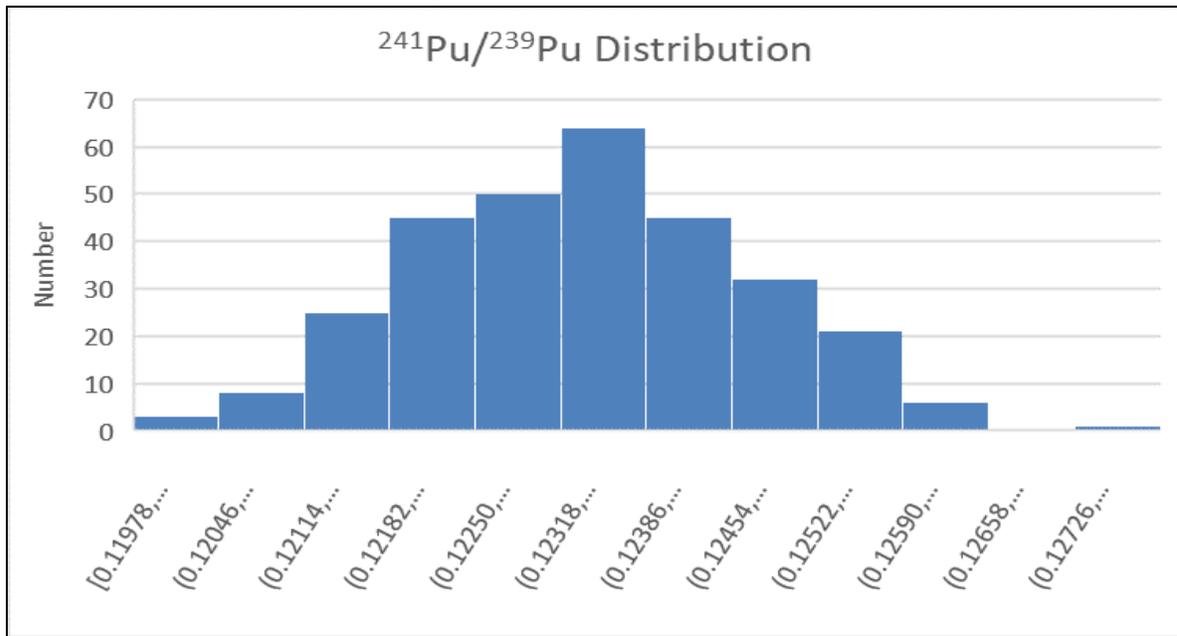


Figure 46: Distribution of values generated for the $^{241}\text{Pu}/^{239}\text{Pu}$ ratio

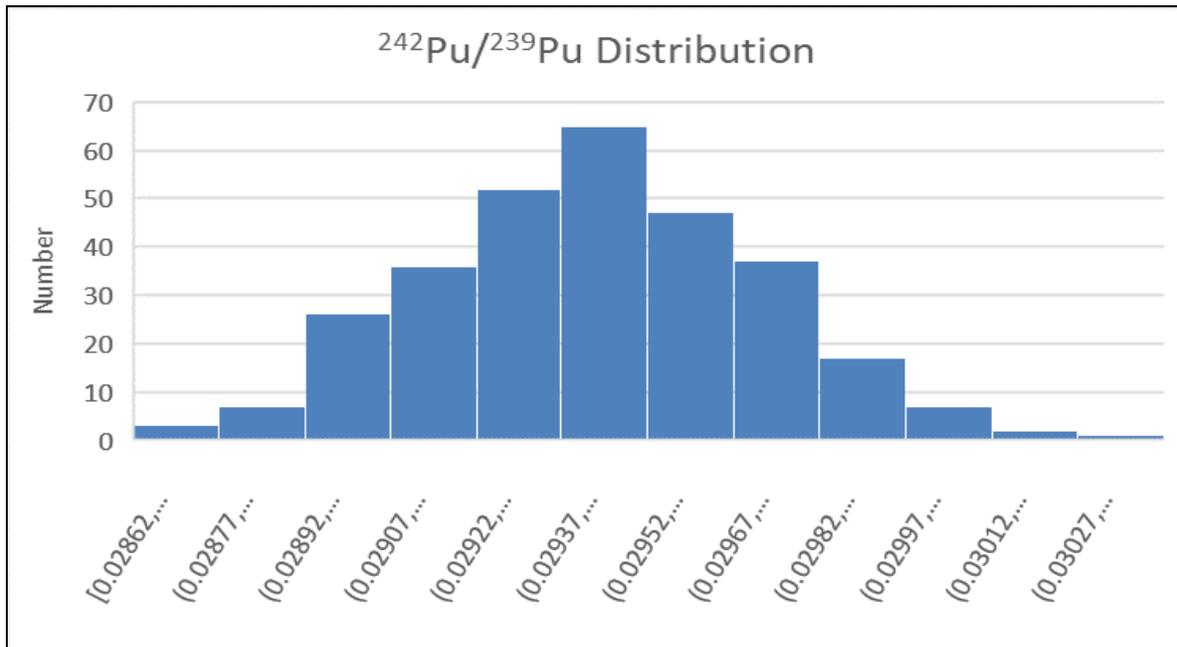


Figure 47: Distribution of values generated for the $^{242}\text{Pu}/^{239}\text{Pu}$ ratio

Table 22 shows the average results of the 300 datapoints generated for the three different uncertainty values chosen. Note that the results change as expected with the probability of attributing the sample to a PWR reactor type decreasing with increasing uncertainty. As the distribution of the data expands, it is expected that the probability of selecting a different reactor type, or the outlier category will increase.

PWR Value with 10% Uncertainty							
	CANDU	BWR	Hanford	MAGNOX	RBMK-1000	PWR	Outlier
Probability	0.00%	6.97%	0.00%	0.00%	0.00%	52.31%	40.71%
Variance	0.00%	0.68%	0.00%	0.00%	0.00%	6.21%	7.60%
PWR Value with 5% Uncertainty							
	CANDU	BWR	Hanford	MAGNOX	RBMK-1000	PWR	Outlier
Probability	0.00%	5.50%	0.00%	0.00%	0.00%	71.75%	22.75%
Variance	0.00%	0.17%	0.00%	0.00%	0.00%	1.60%	2.07%
PWR Value with 1% Uncertainty							
	CANDU	BWR	Hanford	MAGNOX	RBMK-1000	PWR	Outlier
Probability	0.00%	4.28%	0.00%	0.00%	0.00%	79.48%	16.25%
Variance	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.02%

Table 22: Comparison of results for a PWR reactor sample input with different levels of uncertainty applied to the value

This method represents an informal, or ad hoc, approach to properly applying the uncertainty in the sample to be tested to the classification algorithm and shows that this approach shows promise. The results behave as expected with the probability of detecting the correct reactor type decreasing with increasing uncertainty and that probability being applied to other reactor types and the outlier category. A more elegant solution to the classification tool can be created where all the Monte Carlo analysis is completed behind the scenes and can be incorporated into a more user-friendly program.

Conclusions and Future Work

The development of the Classification Algorithm has demonstrated that this approach shows promise as an initial triaging tool for nuclear security, nuclear forensics, and counterproliferation. Using limited measured data and augmenting it with calculated data, there was sufficient variability in the plutonium isotopic ratios to make classification by reactor type plausible. The application of the statistical method will only improve as additional training data is identified and implemented.

As expected, the reactor types that had more data associated with them performed better in the classification algorithm. Additionally, if there was a significant difference in the reactor type: moderator type (light water, heavy water, graphite), fuel enrichment (natural, LEU), irradiation time (including power level, total burnup), or a combination of the three, classification had better results.

Working within reactor types using the same, or similar, parameters (i.e., Light Water Reactors including BWR (GE) and PWR (Westinghouse, VVER-440, VVER-1000) there was a higher probability of misclassification of the reactor type among the similar types. It is expected that if additional reactor data is obtained for all the reactor types, this issue would be reduced.

Even with the informal addition of sample input uncertainty to the front end of the tool, properly identifying and implementing uncertainty in the classification tool will allow the user to understand how to interpret the results presented from the classification algorithm. With respect to uncertainty in general, it will be important to properly quantify and implement uncertainty in the datasets that expand the tool, as well as the sample itself, moving forward.

FUTURE WORK

As previously discussed, the addition of new reactor data to either add additional information to the current reactor types in the algorithm or to expand the number of reactor types would add to the functionality of the classification tool. Additionally, expanding the variables of interest may also provide increased classification probability. SFCOMPO 2.0 has been developed and is designed to facilitate the search and visualization of experimental assay data of used nuclear fuel. It allows the user to access, plot and export isotopic composition data, reactor operational histories and relevant design data relating to used fuel samples (Michel-Sendis F., 2017). SFCOMPO 2.0 will eventually present the user of assay data with a referenced, standardized, cross-checked source of published experimental data for the use of assay data evaluators. Once this work is completed, including this updated dataset will expand the capabilities of the classification tool.

For the purposes of this work, only plutonium isotopic ratios were used in the regression models. Expanding into other actinides (e.g., uranium isotopic ratios) or fission products would give an increased number of vectors to compare against. Adding more variables and vectors may also allow for the classification of reactor types within the similar designs and moderating materials (BWRs vs. PWRs). Ultimately, the goal would be able to identify a specific reactor, based on known operating parameters, however, it is unlikely that sufficient calculated or measured data exists to create a classification algorithm with that much fidelity.

Additionally, age correcting data may become important as the datasets are expanded. As previously noted, the measured data was ingested into the model without age correction, as noted by the flattening of the $^{241}\text{Pu}/^{239}\text{Pu}$ ratio when compared to $^{240}\text{Pu}/^{239}\text{Pu}$ ratio, which may result in classification errors or decreased probability in correctly identifying the reactor type. Depending on the types of samples this tool is expected to

analyze, it may be important to age correct all the data back to reactor discharge. Sensitivity studies to determine the effect of this may need to be completed.

With the addition of new datasets into the classification algorithm, it will be important to make sure uncertainty is properly accounted for in both the training data as well as the sample being tested. The tool can capture the uncertainty information if it is available, but does not apply it to the generation of the covariance matrices. Moving forward, it will be necessary to apply the proper uncertainty or a weighting factor to the measured or calculated data used to generate the parameters for the algorithm. Additionally, a more elegant solution to apply uncertainty in the sample being tested would allow for greater confidence in the algorithm's output. It is understood that incorporating any level of uncertainty to the data will generate a more disperse range of data for each reactor type and will reduce the tool's capability to identify a reactor type of origin, but the user will have additional confidence in the results provided.

Ultimately, this classification algorithm provides a high-level triaging capability to quickly identify or rule out specific reactor types of origin. In the event of interdicted plutonium, there would likely be resource constraints on data analysis. Additional classification algorithms could also be developed that focus on a specific type of reactor (e.g., only focusing on graphite reactors) to further refine the probability to a specific reactor. This tool, and potentially follow-on versions, could help the analysts focus their efforts on more probable reactor types, ultimately reducing the amount of time for attribution of the material.

Appendix A

Appendix A contains the various matrices and parameters that are calculated within the excel workbook for the development and implementation of the Classification Algorithm. These calculations are completed ultimately to produce the information in Table 23, of which the final column shows the posterior probably for the reactor type.

	Ln(det(Ci))	Quadratic.i	lglik.i	Q.i	p(ijY)
CANDU (i=1)	-15.963	5097.821	-2540.929	0.000	0.00%
BWR (i=2)	-8.434	3.588	2.423	1.612	8.18%
Hanford (i=3)	-8.383	33.820	-12.719	0.000	0.00%
MAGNOX (i=4)	-10.168	1588.232	-789.032	0.000	0.00%
RBMK-1000 (i=5)	-14.005	488.112	-237.054	0.000	0.00%
PWR (i=6)	-9.776	0.340	4.718	15.992	81.13%
Outlier (i=7)			2.692	2.108	10.69%

Table 23: Calculation of Posterior Probability for Version 3 of the Classification Algorithm with data from a PWR Type Reactor

For reference, the equations from the main body text are repeated here along with the excel spreadsheet data and tables to help follow the calculations being completed.

Fist we identify the plutonium isotopic ratios and then apply the transformation

$$Y = (Y1, Y2, Y3, Y4)$$

The Transformed Burnup, B' is given by:

$$B' = \log(Y2) = \log(X) = \log\left(\frac{{}^{240}\text{Pu}}{{}^{239}\text{Pu}}\right)$$

and the Transformed Isotopic Ratios are, Y' , are given by:

$$Y' = \log(Y1, Y3, Y4) = \log\left(\frac{{}^{238}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{241}\text{Pu}}{{}^{239}\text{Pu}}, \frac{{}^{242}\text{Pu}}{{}^{239}\text{Pu}}\right)$$

	238/239	240/239	241/239	242/239	
Y	0.0074197	0.216231263	0.108115632	0.015475	B'
Y'	-4.903616625	-1.53140678	-2.224553961	-4.16851	-1.53141

Table 24: Identification of Y, and transformed Y (Y') along with the transformed Burnup (B')

j Isotopic Ratio	α_{1ij}	α_{2ij}	α_{3ij}
CANDU (i=1)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-5.36387	1.218365	-0.07578
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.65312	1.082261	-0.11127
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-1.17076	2.719197	-0.04112
BWR (i=2)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.69421	2.5760113	0.047003
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.11996	0.5956616	-0.2092
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-0.38217	2.6580769	-0.07435
Hanford (i=3)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-5.09915	0.92552	-0.13603
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.10664	1.251023	-0.09743
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-2.08476	2.071997	-0.12125
MAGNOX (i=4)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-3.47827	2.291904	0.050139
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.3521	1.254888	-0.09043
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-2.53673	2.157641	-0.09995
RBMK-1000 (i=5)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-3.22157	2.1961705	0.021492
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-1.21277	0.9672895	-0.13835
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	-1.33958	2.5251331	-0.06752
PWR (i=6)			
$^{238}\text{Pu}/^{239}\text{Pu}$ (j=1)	-1.32174	2.4967816	0.019919
$^{241}\text{Pu}/^{239}\text{Pu}$ (j=2)	-0.67992	0.7397514	-0.19178
$^{242}\text{Pu}/^{239}\text{Pu}$ (j=3)	0.02367	2.6739753	-0.08158

Table 25: Alpha Parameter Estimates from Regression Fit

	$^{238}\text{Pu}/^{239}\text{Pu}$	$^{241}\text{Pu}/^{239}\text{Pu}$	$^{242}\text{Pu}/^{239}\text{Pu}$
CANDU (i=1)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.031827403	-0.029813002	-0.004333687
$^{241}\text{Pu}/^{239}\text{Pu}$	-0.029813002	0.034547026	0.004684451
$^{242}\text{Pu}/^{239}\text{Pu}$	-0.004333687	0.004684451	0.001203521
BWR (i=2)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.19245095	0.01740169	0.07723688
$^{241}\text{Pu}/^{239}\text{Pu}$	0.01740169	0.05526846	0.03107564
$^{242}\text{Pu}/^{239}\text{Pu}$	0.07723688	0.03107564	0.06284208
Hanford (i=3)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.24823887	0.07410551	0.08122217
$^{241}\text{Pu}/^{239}\text{Pu}$	0.07410551	0.03563191	0.03119836
$^{242}\text{Pu}/^{239}\text{Pu}$	0.08122217	0.03119836	0.098336
MAGNOX (i=4)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.05511872	-0.015894612	0.02709316
$^{241}\text{Pu}/^{239}\text{Pu}$	-0.01589461	0.008077273	0.01418814
$^{242}\text{Pu}/^{239}\text{Pu}$	0.02709316	0.01418814	0.35119417
RBMK-1000 (i=5)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.24946013	0.019223538	0.027386961
$^{241}\text{Pu}/^{239}\text{Pu}$	0.01922354	0.002439338	0.003286498
$^{242}\text{Pu}/^{239}\text{Pu}$	0.02738696	0.003286498	0.007913504
PWR (i=6)			
$^{238}\text{Pu}/^{239}\text{Pu}$	0.11676202	0.03306564	0.05713494
$^{241}\text{Pu}/^{239}\text{Pu}$	0.03306564	0.03263444	0.02594396
$^{242}\text{Pu}/^{239}\text{Pu}$	0.05713494	0.02594396	0.05295754

Table 26: Covariance Matrix (C_i) Estimates from Regression Fit

	$^{238}\text{Pu}/^{239}\text{Pu}$	$^{241}\text{Pu}/^{239}\text{Pu}$	$^{242}\text{Pu}/^{239}\text{Pu}$
	CANDU (i=1)		
$^{238}\text{Pu}/^{239}\text{Pu}$	168.0514451	133.3493218	86.09238987
$^{241}\text{Pu}/^{239}\text{Pu}$	133.3493218	167.1109143	-170.2742763
$^{242}\text{Pu}/^{239}\text{Pu}$	86.09238987	-170.2742763	1803.656916
	BWR (i=2)		
$^{238}\text{Pu}/^{239}\text{Pu}$	11.53571057	6.011140365	-17.15064694
$^{241}\text{Pu}/^{239}\text{Pu}$	6.011140365	28.19409471	-21.33015432
$^{242}\text{Pu}/^{239}\text{Pu}$	-17.15064694	-21.33015432	47.53997093
	Hanford (i=3)		
$^{238}\text{Pu}/^{239}\text{Pu}$	11.06689295	-20.78732392	-2.545829051
$^{241}\text{Pu}/^{239}\text{Pu}$	-20.78732392	77.90489692	-7.546711914
$^{242}\text{Pu}/^{239}\text{Pu}$	-2.545829051	-7.546711914	14.66627476
	MAGNOX (i=4)		
$^{238}\text{Pu}/^{239}\text{Pu}$	68.65649697	155.4377484	-11.57620012
$^{241}\text{Pu}/^{239}\text{Pu}$	155.4377301	485.1706556	-31.59212033
$^{242}\text{Pu}/^{239}\text{Pu}$	-11.57619938	-31.59212174	5.016795319
	RBMK-1000 (i=5)		
$^{238}\text{Pu}/^{239}\text{Pu}$	10.27413494	-75.06047414	-4.383802006
$^{241}\text{Pu}/^{239}\text{Pu}$	-75.06049724	1479.085975	-354.4996236
$^{242}\text{Pu}/^{239}\text{Pu}$	-4.383791116	-354.499713	288.7622618
	PWR (i=6)		
$^{238}\text{Pu}/^{239}\text{Pu}$	18.57789565	-4.732168515	-17.72506357
$^{241}\text{Pu}/^{239}\text{Pu}$	-4.732168515	51.39493555	-20.07298655
$^{242}\text{Pu}/^{239}\text{Pu}$	-17.72506357	-20.07298655	47.84008478

Table 27: Inverse Covariance Matrix $(C_i)^{-1}$ Estimates from Regression Fit

Calculate the loglikelihood for each category, i (Values taken from Table 25, Table 26, Table 27):

$$lglik_i = -0.5[\log(|C_i|) + (Y - \mu_i)^T C^{-1}(Y - \mu_i)]$$

With

$$\mu_i = \alpha_{1i} + \alpha_{2i}B'$$

CANDU (i=1)				
Determinate	Inverse	Ln(det(Ci))	ui	Y-ui
1.16833E-07	168.0514451	-15.96251846	-7.40741	2.50379
			-3.57145	1.346895
			-5.43138	1.262875
BWR (i=2)				
Determinate	Inverse	Ln(det(Ci))	ui	Y-ui
0.000217368	11.53571057	-8.433920683	-5.5289	0.625279
			-2.52279	0.298234
			-4.62713	0.458622
Hanford (i=3)				
Determinate	Inverse	Ln(det(Ci))	ui	Y-ui
0.000228661	11.06689295	-8.383271966	-6.83552	1.931906
			-3.25096	1.026404
			-5.54218	1.373675
MAGNOX (i=4)				
Determinate	Inverse	Ln(det(Ci))	ui	Y-ui
3.83851E-05	68.65649697	-10.16784068	-6.87052	1.966905
			-3.48591	1.261354
			-6.07535	1.906848
RBMK-1000 (i=5)				
Determinate	Inverse	Ln(det(Ci))	ui	Y-ui
8.27577E-07	10.27413494	-14.00476319	-6.53439	1.630777
			-3.01854	0.793991
			-5.36492	1.196414
PWR (i=6)				
Determinate	Inverse	Ln(det(Ci))	ui	Y-ui
5.6796E-05	18.57789565	-9.776044241	-5.09862	0.195002
			-2.26254	0.037986
			-4.2626	0.094094

Table 28: Calculation of the Covariance Matrix Determinant, Inverse, and natural log of the Determinant and μ_i and $(Y - \mu_i)$ for the Calculation of the loglikelihood

Calculate the Quadratic as a sub-step for calculating the loglikelihood for each category, i:

$$Quadratic = [(Y - \mu_i)^T C^{-1}(Y - \mu_i)]$$

(C_i)⁻¹(Y-μ_i)			Quadratic
CANDU (i=1)			5097.82147
709.0971	343.9245	2264.009	
BWR (i=2)			3.58762991
1.140093	2.384589	4.717581	
Hanford (i=3)			33.82048466
-3.45313	29.43598	7.482418	
MAGNOX (i=4)			1588.23177
309.0288	857.462	-53.0519	
RBMK-1000 (i=5)			488.112276
-48.0873	627.845	56.86079	
PWR (i=6)			0.340102939
1.775148	-0.85923	0.282537	

Table 29: Calculation of the $C^{-1}(Y - \mu_i)$ Matrix and the *Quadratic* term for the Calculation of the loglikelihood

Finally, calculate Q and the posterior distribution $p(i|Y)$ using the formulas (See Table 23)

$$Q_i = \exp(\lglik_i) \times prior_i$$

And finally

$$p(i|Y) = \frac{Q_i}{\sum_{i=1}^6 Q_i}$$

Appendix B

Appendix B contains the R scripts generated to format the reactor data and perform the various regression fits and overall algorithm development. As development continued throughout the project, the corresponding models and scripts were updated, sometimes resulting in different versions.

DATA INPUT SCRIPTS

The inputDat.r script was used to input the SFCOMPO/NF data and put it in a form that could be used for further analysis. As the analysis progressed, Input.r was developed to incorporate the calculated data into the same form.

InputDat.r

```
#####  
#### Input data from SFCOMPO/NF database (FY11 Summary Report SFCOMPO/NF):  
# read in Data...  
raw.data=read.csv("InputData/DataFrameSheetv2.csv");  
xx=raw.data;  
names(xx)[c(6,8,10,12)]=c("sd.238.239","sd.240.239","sd.241.239","sd.242.239");  
  
# eliminate rows without any data...  
ToRemove=apply(is.na(xx[,7:12]),1,all);  
cat("Note: There are",sum(ToRemove),"blank records. Why is this?\n");  
xx=xx[!ToRemove,-1];  
xx$reactor.types[xx$reactor.types=="VVER-440"]="VVER-440";  
xx$reactor.types=factor(xx$reactor.types);  
xx$reactor=factor(xx$reactor);  
xx$sample.id=factor(xx$sample.id);  
  
# Check out sample ID's; They should be unique;  
sampid=table(xx$sample.id)  
cat("Here are non-unique sample id's:\n");  
print(sampid[sampid>1]);  
  
# check out reactor type/reactor  
print(table(xx$reactor,xx$reactor.type));
```

```

rtbl=cbind(x1=c("BWR","CANDU","Graphite","MAGNOX","PWR","RBMK-
1000","VVER-1000","VVER-440"),
x2=c("BWR","HWR","Graphite","Graphite","PWR","Graphite","PWR","PWR"),
x3=c("GE","CANDU","Hanford","MAGNOX","Westinghouse","RBMK-
1000","VVER-1000","VVER-440"));
id=match(as.character(xx$reactor.types),rtbl[,1]);
### OK Here is my revised dataset..
### include reactor series..
revised.data=data.frame(
  reactor.type=rtbl[id,2],
    reactor.series=rtbl[id,3],
    reactor=xx$reactor,
    sample.id=xx$sample.id,
    Pu.240=log(xx$Pu.240.239),
    Pu.238=log(xx$Pu.238.239),
    Pu.241=log(xx$Pu.241.239),
    Pu.242=log(xx$Pu.242.239),
    sd.240=xx$sd.240.239/xx$Pu.240.239,
    sd.238=xx$sd.238.239/xx$Pu.238.239,
    sd.241=xx$sd.241.239/xx$Pu.241.239,
    sd.242=xx$sd.242.239/xx$Pu.242.239);

### Note: Isotopic ratios that are zero are now represented by -Inf. HOWEVER, people
### constructing this data base may have also used "NA" (i.e. left field blank), so
### it is not clear that one can determine which is which...

#####
## Check data, and if OK move to rdata...
# move.obj(c("revised.data","raw.data"));

#####
# Read in Reactor Physics Model Data...

# read in Hanford data which is in a different format than the rest.
iput <- readLines("InputData/Hanford.txt")
# strip out trailing white space
iput <- sub("\t*$", "", iput);
jput=strsplit(iput,"\t");
Days=as.numeric(jput[[1]][-(1:2)]);
# Burnup in MWd/MT
Burnup=as.numeric(jput[[2]][-(1:2)]);

```

```

Nt=length(Days); Nr=5;
Nloc=sum(sapply(jput,length)==1);
xx=array(NA,dim=c(Nt,Nr,Nloc));
locs=as.character(jput[sapply(jput,length)==1]);
dimnames(xx)=list(round(Burnup,2),
  c("Days",jput[[4]][1],jput[[5]][1],jput[[6]][1],jput[[7]][1]),
  locs);
xx[, "Days",]=Days;
for(i in 3:length(jput)){
pline=jput[[i]];
if(length(pline)==1){
  loca=pline[1]
} else {
  irat=pline[1];
xx[,irat,loca]=as.numeric(pline[-1]);
}
}

## Problem: Are zero Burnup numbers meaningful? They should represent limits, and I
## would expect Pu.xxx/Pu.239=0 as a limit... So I have zeroed out first line...
xx[1,,]=0;
xx=xx[,c(1,2,5,3,4),]
Hanford=xx;

# read in other Reactors....
iput=read.table("InputData/Others.txt",header=F);
for(jj in c("Magnox","Candu","PWR","BWR")){
u=iput[,1]==jj;
xx=t(iput[u,3:12]);
rownames(xx)=NULL;
colnames(xx)=iput[u,2]
### put in "zero burnup" line.
xx=rbind(rep(0,5),xx);
xx=xx[,c(1,2,5,3,4)]
assign(jj,xx)
}

Physics.model.data=list(Hanford=Hanford,MAGNOX=Magnox,CANDU=Candu,Westin
ghouse=PWR,GE=BWR);

#####
### If everything OK, move to rdata...
# move.obj("Physics.model.data");

```

```
#####
#### Combine the two datasets...
## Eliminate SD's
yy=Physics.model.data;
Hanford=yy$Hanford;
yy$Hanford=apply(Hanford,1:2,mean);
for(i in 1:length(yy)) yy[[i]]=yy[[i]][-1,];

ctype=c("Graphite","Graphite","HWR","PWR","BWR");
xx=revised.data[,1:8];
for(i in 1:4) xx[,i]=as.character(xx[,i]);
for(i in 1:length(yy)){
  nobs=dim(yy[[i]])[1];
  zz=data.frame(
    reactor.type=rep(ctype[i],nobs),
    reactor.series=rep(names(yy)[i],nobs),
    reactor=rep("Model",nobs),
    sample.id=round(yy[[i]][,1]),
    Pu.240=log(yy[[i]][,2]),
    Pu.238=log(yy[[i]][,3]),
    Pu.241=log(yy[[i]][,4]),
    Pu.242=log(yy[[i]][,5]),
    stringsAsFactors=F);
  xx=rbind(xx,zz);
}

for(i in 1:3) xx[,i]=factor(xx[,i]);

combined.data=xx
##### Save data...
# move.obj("combined.data");
```

Input.r

```
# Read in Reactor Model Data....
library(openxlsx);

Iput="RD.2016.08.xlsx"
rtype=c("BWR","CANDU","Graphite","Heavy Water","MAGNOX","PWR","RBMK");
ModelRuns=NULL;
for(i in rtype){
```

```

cat(i,"\n");
xx=openxlsx::read.xlsx(Iput,colNames=F,rowNames=T,sheet=i)
print(rownames(xx));
chunk=data.frame(Reactor=i,t(xx),stringsAsFactors=F);
print(chunk[1:3,])
ModelRuns=rbind(ModelRuns,chunk);
}
rownames(ModelRuns)=NULL;

```

```

ModelRuns$Reactor[ModelRuns$Reactor=="Heavy Water"]="HW";

```

```

## add zero burnup points
# Starting Values,
# Yeah, 1 MT (1000 kg, 1E6 g) is the starting condition.
#
# PWR & BWR: (3.5% enriched)
# U-234: 311.5 g
# U-235: 35000 g
# U-236: 161 g
# U-238: 964527.5 g
#
# CANDU, Graphite, MAGNOX: (Natural 0.711%)
# U-234: 63.279 g
# U-235: 7110 g
# U-236: 32.706 g
# U-238: 992794 g
#
# Heavy Water: (1.1% enriched)
# U-234: 97.9 g
# U-235: 11000 g
# U-236: 50.6 g
# U-238: 988851.5 g
#
# RBMK: (2% enriched)
# U-234: 178 g
# U-235: 20000 g
# U-236: 92 g
# U-238: 979730 g

```

```

res=read.table(header=T,as.is=T,text="
Reactor      burnup days  pu238 pu239 pu240 pu241 pu242 u235  u238  total
BWR  0      0      0      0      0      0      3.5+4 964527.5 1e6
PWR  0      0      0      0      0      0      3.5+4 964527.5 1e6

```

```

CANDU      0      0      0      0      0      0      0      0      7110  992794
           1e6
Graphite 0    0      0      0      0      0      0      7110  992794      1e6
MAGNOX     0      0      0      0      0      0      0      0      7110  992794
           1e6
HW         0      0      0      0      0      0      0      11000 988851.5 1e6
RBMK0      0      0      0      0      0      0      0      20000 979730      1e6
");

```

```

xx=rbind(res,ModelRuns);
id=order(xx$Reactor,xx$days)
xx=xx[id,];
ModelRuns=xx;

```

```

library(lattice)
xyplot(pu240/pu239~burnup | Reactor,data=xx)

```

DATA ANALYSIS

This section identifies the R scripts used to analyze the input data and develop the classification algorithms

Initial.Analysis.r

Script to read in and clean up data...

```

raw.data=read.csv("InputData/DataFrameSheetv2.csv");
xx=raw.data;
names(xx)[c(6,8,10,12)]=c("sd.238.239","sd.240.239","sd.241.239","sd.242.239");

```

```

# eliminate rows without any data...
ToRemove=apply(is.na(xx[,7:12]),1,all);
cat("Note: There are",sum(ToRemove),"blank records. Why is this?\n");
xx=xx[!ToRemove,-1];
xx$reactor.types[xx$reactor.types=="VVER-440"]="VVER-440";
xx$reactor.types=factor(xx$reactor.types);
xx$reactor=factor(xx$reactor);
xx$sample.id=factor(xx$sample.id);

```

```

# Check out sample ID's; They should be unique;
sampid=table(xx$sample.id)
cat("Here are non-unique sample id's:\n");

```

```

print(sampid[sampid>1]);

# check out reactor type/reactor
print(table(xx$reactor,xx$reactor.type));

## Do some plots..

bwplot(Pu.240.239~reactor.types,xx,main='Here is "Burnup" Distribution')
xyplot(Pu.241.239~Pu.240.239|reactor.types,xx,main='Check this with spreadsheet')

plot(xx$Pu.240.239,xx$Pu.241.239,pch=as.numeric(xx$reactor.type),
      main="Linear Relationship");
legend("bottomright",levels(xx$reactor.type),pch=(1:length(levels(xx$reactor.type))));

plot(xx$Pu.240.239,xx$Pu.241.239,pch=as.numeric(xx$reactor.type),log="xy",
      main="Log Relationship");
legend("bottomright",levels(xx$reactor.type),pch=(1:length(levels(xx$reactor.type))));

# So it looks like log-log is the way to go
xyplot(log(Pu.238.239)~log(Pu.240.239)|reactor.types,xx,main='Pu 238')
xyplot(log(Pu.241.239)~log(Pu.240.239)|reactor.types,xx,main='Pu 241')
xyplot(log(Pu.242.239)~log(Pu.240.239)|reactor.types,xx,main='Pu 242')

# "reactor" is important;
for(i in levels(xx$reactor.types)){
  u=xx$reactor.types==i;
  y=xx$Pu.241.239[u]; x=xx$Pu.240.239[u]; react=factor(xx$reactor[u]);
  plot(x,y,pch=as.numeric(react),
       main=i,xlab="Pu.240.239",ylab="Pu.241.239");
  legend("bottomright",levels(react),pch=(1:length(levels(react))));
}

rtbl=cbind(x1=c("BWR","CANDU","Graphite","MAGNOX","PWR","RBMK-
1000","VVER-1000","VVER-440"),
           x2=c("BWR","HWR","Graphite","Graphite","PWR","Graphite","PWR","PWR"),
           x3=c("GE","CANDU","Hanford","MAGNOX","Westinghouse","RBMK-
1000","VVER-1000","VVER-440"));
id=match(as.character(xx$reactor.types),rtbl[,1]);
### revised dataset.
### include reactor series.
revised.data=data.frame(
  reactor.type=rtbl[id,2],
  reactor.series=rtbl[id,3],

```

```

reactor=xx$reactor,
sample.id=xx$sample.id,
Pu.240=log(xx$Pu.240.239),
Pu.238=log(xx$Pu.238.239),
Pu.241=log(xx$Pu.241.239),
Pu.242=log(xx$Pu.242.239),
sd.240=xx$sd.240.239/xx$Pu.240.239,
sd.238=xx$sd.238.239/xx$Pu.238.239,
sd.241=xx$sd.241.239/xx$Pu.241.239,
sd.242=xx$sd.242.239/xx$Pu.242.239);

```

Analysis Scheme;

0. log the isotopic ratios. Use reactor.series as discriminant variable.

1. Regress Pu.240.239 against other variables to obtain Beta and residuals...

2. Calculate covariance matrix from the residuals

```
xx=revised.data;
```

```
##### eliminate any data point with missing Pu.240;
```

```
xx=xx[!is.na(xx$Pu.240),];
```

```
### Get rid of log(0);
```

```
xx$Pu.238[is.na(xx$Pu.238) | xx$Pu.238== -Inf]=NA
```

```
xx$Pu.241[is.na(xx$Pu.241) | xx$Pu.241== -Inf]=NA
```

```
xx$Pu.242[is.na(xx$Pu.242) | xx$Pu.242== -Inf]=NA
```

```
##### eliminate any category with less than 10 observations..
```

```
series=levels(xx$reactor.series)[table(xx$reactor.series)>10];
```

```
regfits=vector(length(series),mode="list");
```

```
names(regfits)=series;
```

```
for(ii in series){
```

```
  cat("Working on",ii,"\n");
```

```
  yy=xx[xx$reactor.series==ii,];
```

```
  res238=lm(Pu.238~Pu.240,data=yy)
```

```
  res241=lm(Pu.241~Pu.240,data=yy)
```

```
  res242=lm(Pu.242~Pu.240,data=yy)
```

```
  Beta=rbind(Pu.238=res238$coeff,
```

```
  Pu.241=res241$coeff,
```

```
  Pu.242=res242$coeff);
```

```
  resid=yy[,c("Pu.238","Pu.241","Pu.242")];
```

```
  resid$Pu.238[!is.na(resid$Pu.238)]=res238$resid;
```

```

resid$Pu.241[!is.na(resid$Pu.241)]=res241$resid;
resid$Pu.242[!is.na(resid$Pu.242)]=res242$resid;
Cor=cor(resid,use="complete")
Sigma=c(Pu.238=summary(res238)$sigma,
Pu.241=summary(res241)$sigma,
Pu.242=summary(res242)$sigma);
data.stats=c(Nobs=dim(resid)[1],Ncomplete=sum(apply(!is.na(resid),1,all)));
regfits[[ii]]=list(Beta=Beta,Cor=Cor,Sigma=Sigma, Nobs=data.stats,
Pu.240.quant=quantile(yy$Pu.240),
res238=res238, res241=res241,res242=res242,data=yy);
}
##### 3. Construct Discriminant function. Disc(pnts,prior=,cov=,beta=);
loglikmvn=function(xpnts,Mu,Cov)
sqrtdet= -0.5*log(eigen(cov)$values)
zz=solve(Cov,xpnts-Mu)
qterm = -0.5*apply(xpnts*zz,1,sum)
loglik= sqrtdet +

D.bayes=function(xpnts,prior,Beta,Cov){
# This function classifies points in table "xpnts";
names(pnst)=c(Pu.240,Pu.238,Pu.241,Pu.242);
# Classifier is multivariate normal with mean mu=Beta %*% Pu.240, and covariance
defined by
# "cov".

groups=names(prior)
##### 4. Calculate confusion matrix by Monte Carlo....
MonteCarlo(prior=,burnup=, cov= ,beta=);

```

CrossVal2.r

```

#####
#####
fitDbayes=function(xx,linear=T){
# This fits quadratic model with weights. The weights distinguish between
# real data wt=1, and calculated data wt=Var(data)/Var(Calc). So if Calculated
# are more precise, then wt>1, but if you believe there are "biases" in the calculation
# the wt<1.
# if linear=T, then y=B1+B2*X, otherwise y=B1+B2*X+B3*X^2

# add weight vector if not present...
if(is.null(xx$wt)) xx$wt=1;

```

```

groups=levels(xx$reactor.series);
Beta=array(0,dim=c(3,3,length(groups)));
dimnames(Beta)=list(c("Pu.238","Pu.241","Pu.242"),
  c("Beta.1","Beta.2","Beta.3"),groups);
Cov=array(0,dim=c(3,3,length(groups)));
dimnames(Cov)=list(c("Pu.238","Pu.241","Pu.242"),
  c("Pu.238","Pu.241","Pu.242"),groups);
regfits=vector(length=length(groups),mode="list");
names(regfits)=groups;
for(ii in groups){
  u=xx$reactor.series==ii;
  if(linear){
    res238=lm(Pu.238~Pu.240,weights=wt,data=xx,subset=u)
    res241=lm(Pu.241~Pu.240,weights=wt,data=xx,subset=u)
    res242=lm(Pu.242~Pu.240,weights=wt,data=xx,subset=u)
    Beta[,1:2,ii]=rbind(Pu.238=res238$coeff,
      Pu.241=res241$coeff,
      Pu.242=res242$coeff);
  }else{
    res238=lm(Pu.238~Pu.240+I(Pu.240^2),weights=wt,data=xx,subset=u)
    res241=lm(Pu.241~Pu.240+I(Pu.240^2),weights=wt,data=xx,subset=u)
    res242=lm(Pu.242~Pu.240+I(Pu.240^2),weights=wt,data=xx,subset=u)
    Beta[,ii]=rbind(Pu.238=res238$coeff,
      Pu.241=res241$coeff,
      Pu.242=res242$coeff);
  }
  resid=xx[u,c("Pu.238","Pu.241","Pu.242")];
  resid$Pu.238[!is.na(resid$Pu.238)]=res238$resid;
  resid$Pu.241[!is.na(resid$Pu.241)]=res241$resid;
  resid$Pu.242[!is.na(resid$Pu.242)]=res242$resid;
  Cor=cor(resid,use="complete")
  Sigma=c(Pu.238=summary(res238)$sigma,
    Pu.241=summary(res241)$sigma,
    Pu.242=summary(res242)$sigma);
  Cov[,ii]=Cor*outer(Sigma,Sigma);
  data.stats=c(Nobs=dim(resid)[1],Ncomplete=sum(apply(!is.na(resid),1,all)));
  regfits[[ii]]=list(Beta=Beta[,ii],Cor=Cor,Sigma=Sigma, Nobs=data.stats,
    Pu.240.quant=quantile(xx$Pu.240[u]),
    res238=res238, res241=res241,res242=res242);
}
xfit=list(Beta=Beta,Cov=Cov,regfits=regfits);
return(xfit)
}

```

```
#####

#####
plot.fitDbayes=function(xfit,VarMod=c(Sdet=-Inf,Smax=5)){
# plot regression fits and Data...
oldpar=par(mfcol=c(3,1));
Sdet=VarMod[1]; Smax=VarMod[2];
xomax=NULL;
for(i in 1:length(xfit$regfits)){
glab=names(xfit$regfits)[i];

afit=xfit$regfits[[i]]$res238; Beta=xfit$regfits[[i]]$Beta[1,];

plot(afit$model[,2],afit$model[,1],main=glab,xlab="Pu.240/Pu.239",ylab="Pu.238/Pu.23
9");
xo=range(afit$model[,2]); zz=seq(from=xo[1],to=xo[2],length=100);
yp=Beta[1]+Beta[2]*zz + Beta[3]*zz^2; sigma=summary(afit)$sigma;
matlines(zz,outer(yp,c(0,-2*sigma,2*sigma),"+"),col=c(1,2,2),lty=c(1,2,2));
xomax=range(c(xomax,xo));

afit=xfit$regfits[[i]]$res241; Beta=xfit$regfits[[i]]$Beta[2,];

plot(afit$model[,2],afit$model[,1],main=glab,xlab="Pu.240/Pu.239",ylab="Pu.241/Pu.23
9");
xo=range(afit$model[,2]); zz=seq(from=xo[1],to=xo[2],length=100);
yp=Beta[1]+Beta[2]*zz + Beta[3]*zz^2; sigma=summary(afit)$sigma;
matlines(zz,outer(yp,c(0,-2*sigma,2*sigma),"+"),col=c(1,2,2),lty=c(1,2,2));

afit=xfit$regfits[[i]]$res242; Beta=xfit$regfits[[i]]$Beta[3,];

plot(afit$model[,2],afit$model[,1],main=glab,xlab="Pu.240/Pu.239",ylab="Pu.242/Pu.23
9");
xo=range(afit$model[,2]); zz=seq(from=xo[1],to=xo[2],length=100);
yp=Beta[1]+Beta[2]*zz + Beta[3]*zz^2; sigma=summary(afit)$sigma;
matlines(zz,outer(yp,c(0,-2*sigma,2*sigma),"+"),col=c(1,2,2),lty=c(1,2,2));

}

### Plot regression fits for all groups to see discrimination capability...
glab=names(xfit$regfits); ng=length(glab);

zz=seq(from=xomax[1],to=xomax[2],length=100);
isoname=dimnames(xfit$Beta)[[1]]
```

```

for(i in 1:length(isoname)){
yp=cbind(1,zz,zz^2) %*% xfit$Beta[i,];
Sp=sqrt(xfit$Cov[i,i]);
St=pmin(sqrt(Sp^2 + exp(2*(Sdet-yp))),Smax);
matplot(zz,yp, type="l",main=isoname[i],
xlab="Pu.240/Pu.239",ylab=isoname[i], col=1:ng,lty=1:ng,lwd=3);
# plot out error bars....
j1=seq(dim(yp)[1]) %% dim(yp)[2];
for(ij in 1:dim(yp)[2]){
u=j1==(ij-1);
plot.error.bars(zz[u],yp[u,ij],yp[u,ij]-St[u,ij],yp[u,ij]+St[u,ij]);
}
legend("topleft",glab,col=1:ng,lty=1:ng,lwd=3);
}
}
par(oldpar)
}
#####

#####
Dbayes=function(xpnts,xfit,prior=NULL,VarMod=c(Sdet=-Inf,Smax=5),classify=F){
# This function classifies points in table "xpnts";
## I have sped up things by altering form of measurement error..
## Ct = Cr*outer(St,St) with St = Sp + sqrt(Co)/exp(Mu);
# names(xpnts)=c(Pu.240,Pu.238,Pu.241,Pu.242);
# Classifier is multivariate normal with parameters calculated from xfit,
# prior, and VarMod.
# xfit: contains regression model fit. xfit$Cov proportional variability
# VarMod: Measurement error model Sdet=log(Stdev) at detection. Smax=maximum
# stdev.

Beta=xfit$Beta; Cov=xfit$Cov;
groups=dimnames(Beta)[[3]]
if(is.null(prior)){
prior=rep(1,length(groups));
names(prior)=groups;
prior=prior/sum(prior);
}
Sdet=VarMod[1]; Smax=VarMod[2];
# Returns the Posterior for each data-point in xpnts....
if(abs(sum(prior)-1)>1e-6) stop("Prior does not add to one")
# if(any(is.na(xpnts[,c("Pu.238","Pu.240","Pu.241","Pu.242")]))
# stop("NA not allowed in Data");
x=xpnts[,"Pu.240"]

```

```

y=xpnts[,c("Pu.238", "Pu.241", "Pu.242")];
lglik=array(0,dim=c(length(x),length(prior)));
colnames(lglik)=groups;
for(ii in groups){
  Mu= Beta[,ii] %*% rbind(1,x,x^2);
  Sp=sqrt(diag(Cov[,ii]));
  St=pmin(sqrt(Sp^2 + exp(2*(Sdet-Mu))),Smax);
  Cr= Cov[,ii]*outer(1/Sp,1/Sp)
  logDCr=sum(log(eigen(Cr)$values));
  sqlogdet= -0.5*logDCr - apply(log(St),2,sum);
  y1=(t(y)-Mu)/St;
  qterm = -0.5*apply(y1*solve(Cr,y1),2,sum);
  lglik[,ii]= sqlogdet + qterm + log(prior[ii]);
}
post=exp(lglik);
post=post/apply(post,1,sum);
# if one is required to choose a class
if(classify){
  has.na=apply(is.na(post),1,any);
  zz=apply(-post,1,order)[1,];
  post[]=0; post[cbind(1:length(zz),zz)]=1;
  post[has.na,]=NA;
}
return(post)
}

#####
predfun=function(Xtrain,Xtest,prior=NULL,
  VarMod=c(Sdet=-Inf,Smax=5),linear=T,classify=F){
# Xtrain: dataset or fitted results that will be used to construct classifier
# Xtest: dataset that will be used to test classifier

if(all(names(Xtrain)[1:3]==c("Beta", "Cov", "regfits"))) xfit=Xtrain
else xfit=fitDbayes(Xtrain,linear=linear);
# This coding allows use an existing fit...
post=Dbayes(Xtest,xfit,prior=prior,VarMod=VarMod,classify=classify);
confmat=array(NA,dim=rep(dim(post)[2],2));
dimnames(confmat)=list(Actual=dimnames(post)[[2]],
  Pred=dimnames(post)[[2]]);
for(i in 1:dim(post)[2]){
# Note: this returns "NaN" when there isn't anything in a category
confmat[,i]=tapply(post[,i],Xtest$reactor.series,mean,na.rm=T)
}
}

```

```

confmat[is.nan(confmat)]=NA;
return(confmat);

}
#####

#####

Crossval=function(data,pfun=predfun,cv.sets=lapply(1:dim(data)[1],max),
prior=NULL,VarMod=c(Sdet=-Inf,Smax=5),linear=T,classify=F){
# pfun fits the model and then applies it to the cross validation sets
# data contains the data pfun uses. data will be split into a training
# set and test set
# val.sets: is a list with each component containing a set of indices to
# be in a test (cross validation) set. So default is leave one out CV.

res0=pfun(data,data,prior=prior,VarMod=VarMod,linear=linear,classify=classify);
res1=array(0,dim=c(dim(res0),length(cv.sets)));
dimnames(res1)=list(rownames(res0),colnames(res0),names(cv.sets));
for(i in 1:length(cv.sets)){
cat("Working on ",i,"\n");
res1[,i]=pfun(data[-cv.sets[[i]],],data[cv.sets[[i]],],
prior=prior,VarMod=VarMod,linear=linear,classify=classify);
}
av.cm=apply(res1,1:2,mean,na.rm=T);
sd.cm=apply(res1,1:2,sd,na.rm=T);
return(list(cv.sets=cv.sets,cv.runs=res1,Alldat.cm=res0,cv.cm=av.cm,cv.sd=sd.cm));
}
#####

#####

select.sets=function(data,ratio=.10,Nrep=100){
# dataset to use in cross validation exercise
# ratio Proportion of each group that is to be placed in the test data..
Rseed=.Random.seed;
y=data$reactor.series;
Ns=round(pmax(table(y)*ratio,1));
ygrouped=vector(length(levels(y)),mode="list");
names(ygrouped)=levels(y);
for(i in names(ygrouped)) ygrouped[[i]]=seq(y)[y==i];
cvset=vector(Nrep,mode="list");
for(j in 1:length(ygrouped)){
for(i in 1:Nrep) cvset[[i]]= c(cvset[[i]],sample(ygrouped[[j]],Ns[j]));
}
}

```

```

attr(cvset,"random.seed")=Rseed;
return(cvset);
}
#####

#####

SelectByReactor=function(data){
# dataset to use in cross validation exercise
# ratio Proportion of each reactor that is to be placed in the test data..
y=data$reactor;
Ns=table(y);
ygrouped=vector(length(levels(y)),mode="list");
names(ygrouped)=levels(y);
for(i in names(ygrouped)) ygrouped[[i]]=seq(y)[y==i];
cvset=ygrouped;
return(cvset);
}
#####

#####

SimData=function(xfit,Pu.240=c(-5,0),VarMod=c(Sdet=-Inf,Smax=5),nmc=1000){
# this produces an "xdata" set of simulated values.....
Sdet=VarMod[1]; Smax=VarMod[2];
grps=dimnames(xfit$Beta)[[3]]; ng=length(grps);
np=length(Pu.240);
iseed=.Random.seed;
nobs=nmc*ng*length(Pu.240);
rseries=numeric(nobs);
Xtest=array(0,dim=c(nobs,4));
colnames(Xtest)=c("Pu.240","Pu.238","Pu.241","Pu.242");
icnt=0;
idx=cbind(rep(1:ng,np),rep(1:np,rep(ng,np)));
for(i in 1:(ng*np)){
jg=idx[i,1];
jp=idx[i,2];
kslc=icnt+1:nmc;
rseries[kslc]=jg;
Xtest[kslc,1]=Pu.240[jp];
mu = as.vector(xfit$Beta[,jg] %*% c(1,Pu.240[jp],Pu.240[jp]^2));
Sp=sqrt(diag(xfit$Cov[,jg]))
Crl= xfit$Cov[,jg]*outer(1/Sp,1/Sp);
St=pmin(sqrt(Sp^2 + exp(2*(Sdet-mu))),Smax);
Xtest[kslc,2:4]=mvrnorm(n=nmc,rep(0,length(mu)),Crl);

```

```

for(kk in 2:4) Xtest[kslc,kk]=mu[kk-1]+Xtest[kslc,kk]*St[kk-1];
icnt=icnt+nmc;
}
return(data.frame(
  reactor.series=factor(rseries,levels=1:ng,labels=grps),
  Xtest));
}

#####
IntDbayes=function(xfit,Pu.240=seq(from=-5,to=0,length=6),
  prior=NULL,VarMod=c(-Inf,5),nmc=1000,linear=T,classify=F,output=F){
# calculate confusion matrix for each Pu.240 value using MC Integration.
# xfit; Regression Model...
iseed=.Random.seed;
grps=dimnames(xfit$Beta)[[3]]; ng=length(grps);
confmat=array(NA,dim=c(ng,ng,length(Pu.240)));
dimnames(confmat)=list(Actual=grps,Pred=grps,Pu.240=round(Pu.240,2));
mcdata=vector(length(Pu.240),mode="list");
for(i in 1:length(Pu.240)){
xsim=SimData(xfit,Pu.240=Pu.240[i],VarMod=VarMod,nmc=nmc);
confmat[,i]=predfun(xfit,xsim,
  prior=prior,classify=classify,VarMod=VarMod,linear=linear);
if(output) mcdata[[i]]=xsim
}
res1=list(xfit=xfit,prior=prior,VarMod=VarMod,classify=classify,nmc=nmc,iseed=iseed);
res=list(Pu.240=Pu.240,confmat=confmat,mcdata=mcdata,Arguments=res1);
return(res);
}

#####
plot.IntDbayes=function(res){
Pu.240=exp(res$Pu.240);
error=res$confmat[1,];
nx=dim(error)[1];
for(i in 1:nx) error[i,]=1-res$confmat[i,i];
# error=cbind(t(error),total=apply(error,2,mean));
error=t(error);
matplot(Pu.240,error,type="l",log="x",xlab="Pu.240/Pu.239",ylab="Error",
  main="Classification Error",lty=1:nx,col=1:nx,lwd=3);
legend("topright",colnames(error),lty=1:nx,col=1:nx,lwd=3);
invisible();
}

```

ANALYSIS SCRIPTS

This final script is used to develop the Version 3 of the classification algorithm and includes outlier data. The data generated from this script is used in the classification tool.

Fun.r

```
#####  
Make.outliers=function(xpnts,draws=5,nsiz=3){  
# Randomly generate outlier points...  
# This makes draws*(groups+1) "outlier" data points from the dataset "xpnts".  
# Most of the outliers are produced by averaging nsiz data-points from xpnts  
# together. The last set of outliers are sampled from a uniform multivariate  
# distribution.  
# arguments  
# xpnts: Our standard dataset containing the 4 isotopic ratios used for discrimination  
# draws: The number of random draws to be made from each reactor group  
# nsiz: the number of data points to be averaged together  
# Value: xpnts, with the draws*(groups+1) outlier data points appended to the end.  
# xpnts$reactor.series will have the group "Outlier" added.  
  
groups=levels(xpnts$reactor.series);  
levels(xpnts$reactor.series)=c(levels(xpnts$reactor.series),"Outlier");  
id=1:dim(xpnts)[1];  
res=NULL;  
# This loop averages nsiz=3 randomly chosen data.pnts together to form an "outlier"  
for(ii in groups){  
id1=id[xpnts$reactor.series==ii];  
for(j in 1:draws){  
id2=sample(id1,size=nsiz);  
zz=xpnts[id2[1],]; zz$reactor.series[1]="Outlier";  
zz[1,5:8]=log(apply(exp(xpnts[id2,5:8]),2,mean,na.rm=T));  
res=rbind(res,zz)  
}  
}  
# An alternative; choose values from a uniform distribution that extends over the range  
# of the data.  
n1=draws  
id2=sample(id,size=n1,replace=T);  
# Note that other fields besides the ratios and reactor.series are meaningless for  
# the outlier data points....  
zz=xpnts[id2,]; zz$reactor.series[]="Outlier";
```

```
zz$Pu.240[]=runif(n1,min=min(xpnts$Pu.240,na.rm=T),max=max(xpnts$Pu.240,na.rm=
T));
```

```
zz$Pu.238[]=runif(n1,min=min(xpnts$Pu.238,na.rm=T),max=max(xpnts$Pu.238,na.rm=
T));
```

```
zz$Pu.241[]=runif(n1,min=min(xpnts$Pu.241,na.rm=T),max=max(xpnts$Pu.241,na.rm=
T));
```

```
zz$Pu.242[]=runif(n1,min=min(xpnts$Pu.242,na.rm=T),max=max(xpnts$Pu.242,na.rm=
T));
res=rbind(res,zz)
return(rbind(xpnts,res))
}
```

```
#####
```

```
Dbayes2=function(xpnts,xfit,prior=NULL,VarMod=c(Sdet=-Inf,Smax=5),classify=F){
```

```
# This function classifies points in table "xpnts";
```

```
# MODIFICATION to include "Outlier Category". Outlier category/group is added to
# existing groups.
```

```
# Arguments;
```

```
# xpnts: a standard dataset of ratios. These are the observations to be classified.
```

```
# (i.e. this is the test set).
```

```
# xfit: a regression fit to the data, produced by fitDbayes(); This contains model
# parameters.
```

```
# prior: This is prior on groups, including "Outlier". If missing, categories are
# equally weighted in the prior, except for "Outlier", which is assumed to
# be 1% of population.
```

```
# VarMod: Variance model, which is defined by
```

```
# St=pmin(sqrt(Sp^2 + exp(2*(Sdet-Mu))),Smax);
```

```
# classify: if T, the output identifies a single group
```

```
# be given by qchisq(.99,dof=3)=10. dof is the dimension of the mult.
```

```
# distribution.
```

```
# Value; A matrix of posterior distributions if classify=F, otherwise an indicator vector.
```

```
Beta=xfit$Beta; Cov=xfit$Cov;
```

```
groups=dimnames(Beta)[[3]]
```

```
Sdet=VarMod[1]; Smax=VarMod[2];
```

```
# Returns the Posterior for each data-point in xpnts....
```

```
# if(abs(sum(prior)-1)>1e-6) stop("Prior does not add to one")
```

```
# This works OK for now
```

```
# if(any(is.na(xpnts[,c("Pu.238","Pu.240","Pu.241","Pu.242")])))
```

```

# stop("NA not allowed in Data");
x=xpnts["Pu.240"]
y=xpnts[,c("Pu.238","Pu.241","Pu.242")];
lglik=array(0,dim=c(length(x),length(groups)));
colnames(lglik)=groups;
sqlogdet=lglik;
for(ii in groups){
  Mu= Beta[,ii] %*% rbind(1,x,x^2);
  Sp=sqrt(diag(Cov[,ii]));
  St=pmin(sqrt(Sp^2 + exp(2*(Sdet-Mu))),Smax);
  Cr= Cov[,ii]*outer(1/Sp,1/Sp)
  logDCr=sum(log(eigen(Cr)$values));
  sqlogdet[,ii]= -0.5*logDCr - apply(log(St),2,sum);
  y1=(t(y)-Mu)/St;
  qterm = -0.5*apply(y1*solve(Cr,y1),2,sum);
  lglik[,ii]= sqlogdet[,ii] + qterm;
}

##### Construct prior if not present...
##### Outlier component is set at default of 1%
if(is.null(prior)){
  prout=.01;
  prior=rep(1,dim(lglik)[2]);
  names(prior)=dimnames(lglik)[[2]];
  prior=prior/sum(prior);
  prior=c((1-prout)*prior,Outlier=prout);
}

#####
# add on "Outlier category. log-likelihood associated with
# outliers in a uniform distribution, with a magnitude
# of;
# min.ii(sqlogdet[ii]) + E(qterm)
# where E(qterm) = -0.5*p, p=3 is the dimension of the quadratic term.
# could use max, min, or mean in formula. Min assures that
# the category with highest variability is not overwhelmed by
# outlier category;

P=dim(Cov)[1];
lglik.outlier=apply(sqlogdet,1,min)-0.5*P;
lglik=cbind(lglik,Outlier=lglik.outlier);

post=exp(lglik)*outer(rep(1,dim(lglik)[1]),prior);

```

```

post=post/apply(post,1,sum);
# if one is required to choose a class
if(classify){
  has.na=apply(is.na(post),1,any);
  zz=apply(-post,1,order)[1,];
  post[]=0; post[cbind(1:length(zz),zz)]=1;
  post[has.na,]=NA;
}
return(post)
} # End of Dbayes2

```

```

#####
predfun2=function(Xtrain,Xtest,prior=NULL,
  VarMod=c(Sdet=-Inf,Smax=5),linear=T,classify=F){
# Xtrain: dataset or fitted results that will be used to construct classifier
# Xtest: dataset that will be used to test classifier

if(all(names(Xtrain)[1:3]==c("Beta","Cov","regfits"))) xfit=Xtrain
else{
  u=Xtrain$reactor.series!="Outlier";
  xfit=fitDbayes(Xtrain[u,],linear=linear);
}
# This coding allows me to use an existing fit...
post=Dbayes2(Xtest,xfit,prior=prior,VarMod=VarMod,classify=classify);
confmat=array(NA,dim=rep(dim(post)[2],2));
dimnames(confmat)=list(Actual=dimnames(post)[[2]],
Pred=dimnames(post)[[2]]);
for(i in 1:dim(post)[2]){
# Note: this returns "NaN" when there isn't anything in a category
# Force categories to be correct, even when Outlier is missing...
confmat[,i]=tapply(post[,i],
factor(Xtest$reactor.series,levels=c(dimnames(xfit$Beta)[[3]],"Outlier")),
  mean,na.rm=T)
}
confmat[is.nan(confmat)]=NA;
return(confmat);
}

```

```

#####

#####
IntDbayes2=function(xfit,Pu.240=seq(from=-5,to=0,length=6),
prior=NULL,VarMod=c(-Inf,5),nmc=1000,linear=T,classify=F,output=F){

```

```

# calculate confusion matrix for each Pu.240 value using MC Integration.
# xfit; Regression Model...
iseed=.Random.seed;
grps=dimnames(xfit$Beta)[[3]]; ng=length(grps);
#### only dimension of confmat is changed from previous version...
confmat=array(NA,dim=c(ng+1,ng+1,length(Pu.240)));
dimnames(confmat)=list(Actual=c(grps,"Outlier"),Pred=c(grps,"Outlier"),Pu.240=round(
Pu.240,2));
mcdata=vector(length(Pu.240),mode="list");
for(i in 1:length(Pu.240)){
#####
xsim=SimData(xfit,Pu.240=Pu.240[i],VarMod=VarMod,nmc=nmc);
confmat[,i]=predfun2(xfit,xsim,
prior=prior,classify=classify,VarMod=VarMod,linear=linear);
if(output) mcdata[[i]]=xsim
}
res1=list(xfit=xfit,prior=prior,VarMod=VarMod,classify=classify,nmc=nmc,iseed=iseed);
res=list(Pu.240=Pu.240,confmat=confmat,mcdata=mcdata,Arguments=res1);
return(res);
}

```

```

#####
#####
# The following subroutines are the model without the Outlier category...
#####
#####

```

```

#####
fitDbayes=function(xx,linear=T){
# This fits quadratic model with weights. The weights distinguish between
# real data wt=1, and calculated data wt=Var(data)/Var(Calc). So if Calculated
# are move precise, then wt>1, but if you believe there are "biases" in the calculation
# the wt<1.
# if linear=T, then y=B1+B2*X, otherwise y=B1+B2*X+B3*X^2

```

```

# add weight vector if not present...
if(is.null(xx$wt)) xx$wt=1;
groups=levels(xx$reactor.series);
Beta=array(0,dim=c(3,3,length(groups)));
dimnames(Beta)=list(c("Pu.238","Pu.241","Pu.242"),
c("Beta.1","Beta.2","Beta.3"),groups);
Cov=array(0,dim=c(3,3,length(groups)));
dimnames(Cov)=list(c("Pu.238","Pu.241","Pu.242"),

```

```

c("Pu.238","Pu.241","Pu.242"),groups);
regfits=vector(length=length(groups),mode="list");
names(regfits)=groups;
for(ii in groups){
  u=xx$reactor.series==ii;
  if(linear){
    res238=lm(Pu.238~Pu.240,weights=wt,data=xx,subset=u)
    res241=lm(Pu.241~Pu.240,weights=wt,data=xx,subset=u)
    res242=lm(Pu.242~Pu.240,weights=wt,data=xx,subset=u)
    Beta[,1:2,ii]=rbind(Pu.238=res238$coeff,
    Pu.241=res241$coeff,
      Pu.242=res242$coeff);
  }else{
    res238=lm(Pu.238~Pu.240+I(Pu.240^2),weights=wt,data=xx,subset=u)
    res241=lm(Pu.241~Pu.240+I(Pu.240^2),weights=wt,data=xx,subset=u)
    res242=lm(Pu.242~Pu.240+I(Pu.240^2),weights=wt,data=xx,subset=u)
    Beta[,,ii]=rbind(Pu.238=res238$coeff,
    Pu.241=res241$coeff,
      Pu.242=res242$coeff);
  }
  resid=xx[u,c("Pu.238","Pu.241","Pu.242")];
  resid$Pu.238[!is.na(resid$Pu.238)]=res238$resid;
  resid$Pu.241[!is.na(resid$Pu.241)]=res241$resid;
  resid$Pu.242[!is.na(resid$Pu.242)]=res242$resid;
  Cor=cor(resid,use="complete")
  Sigma=c(Pu.238=summary(res238)$sigma,
    Pu.241=summary(res241)$sigma,
    Pu.242=summary(res242)$sigma);
  Cov[,,ii]=Cor*outer(Sigma,Sigma);
  data.stats=c(Nobs=dim(resid)[1],Ncomplete=sum(apply(!is.na(resid),1,all)));
  regfits[[ii]]=list(Beta=Beta[,,ii],Cor=Cor,Sigma=Sigma, Nobs=data.stats,
  Pu.240.quant=quantile(xx$Pu.240[u]),
  res238=res238, res241=res241,res242=res242);
}
xfit=list(Beta=Beta,Cov=Cov,regfits=regfits);
return(xfit)
}
#####

#####

plot.fitDbayes=function(xfit,VarMod=c(Sdet=-Inf,Smax=5)){
# plot regression fits and Data...
oldpar=par(mfcol=c(3,1));

```

```

Sdet=VarMod[1]; Smax=VarMod[2];
xomax=NULL;
for(i in 1:length(xfit$regfits)){
  glab=names(xfit$regfits)[i];

  afit=xfit$regfits[[i]]$res238; Beta=xfit$regfits[[i]]$Beta[1,];

  plot(afit$model[,2],afit$model[,1],main=glab,xlab="Pu.240/Pu.239",ylab="Pu.238/Pu.23
9");
  xo=range(afit$model[,2]); zz=seq(from=xo[1],to=xo[2],length=100);
  yp=Beta[1]+Beta[2]*zz + Beta[3]*zz^2; sigma=summary(afit)$sigma;
  matlines(zz,outer(yp,c(0,-2*sigma,2*sigma),"+"),col=c(1,2,2),lty=c(1,2,2));
  xomax=range(c(xomax,xo));

  afit=xfit$regfits[[i]]$res241; Beta=xfit$regfits[[i]]$Beta[2,];

  plot(afit$model[,2],afit$model[,1],main=glab,xlab="Pu.240/Pu.239",ylab="Pu.241/Pu.23
9");
  xo=range(afit$model[,2]); zz=seq(from=xo[1],to=xo[2],length=100);
  yp=Beta[1]+Beta[2]*zz + Beta[3]*zz^2; sigma=summary(afit)$sigma;
  matlines(zz,outer(yp,c(0,-2*sigma,2*sigma),"+"),col=c(1,2,2),lty=c(1,2,2));

  afit=xfit$regfits[[i]]$res242; Beta=xfit$regfits[[i]]$Beta[3,];

  plot(afit$model[,2],afit$model[,1],main=glab,xlab="Pu.240/Pu.239",ylab="Pu.242/Pu.23
9");
  xo=range(afit$model[,2]); zz=seq(from=xo[1],to=xo[2],length=100);
  yp=Beta[1]+Beta[2]*zz + Beta[3]*zz^2; sigma=summary(afit)$sigma;
  matlines(zz,outer(yp,c(0,-2*sigma,2*sigma),"+"),col=c(1,2,2),lty=c(1,2,2));

}

### Plot regression fits for all groups to see discrimination capability...
glab=names(xfit$regfits); ng=length(glab);

zz=seq(from=xomax[1],to=xomax[2],length=100);
isoname=dimnames(xfit$Beta)[[1]]
for(i in 1:length(isoname)){
  yp=cbind(1,zz,zz^2) %*% xfit$Beta[i,];
  Sp=sqrt(xfit$Cov[i,i]);
  St=pmin(sqrt(Sp^2 + exp(2*(Sdet-yp))),Smax);
  matplot(zz,yp, type="l",main=isoname[i],
  xlab="Pu.240/Pu.239",ylab=isoname[i], col=1:ng,lty=1:ng,lwd=3);

```

```

# plot out error bars...
j1=seq(dim(yp)[1]) %% dim(yp)[2];
for(ij in 1:dim(yp)[2]){
u=j1==(ij-1);
plot.error.bars(zz[u],yp[u,ij],yp[u,ij]-St[u,ij],yp[u,ij]+St[u,ij]);
}
legend("topleft",glab,col=1:ng,lty=1:ng,lwd=3);
}
par(oldpar)
}
#####

#####
Dbayes=function(xpnts,xfit,prior=NULL,VarMod=c(Sdet=-Inf,Smax=5),classify=F){
# This function classifies points in table "xpnts";
## I have sped up things by altering form of measurement error..
## Ct = Cr*outer(St,St) with St = Sp + sqrt(Co)/exp(Mu);
# names(xpnts)=c(Pu.240,Pu.238,Pu.241,Pu.242);
# Classifier is multivariate normal with parameters calculated from xfit,
# prior, and VarMod.
# xfit: contains regression model fit. xfit$Cov proportional variability
# VarMod: Measurement error model Sdet=log(Stdev) at detection. Smax=maximum
# stdev.

Beta=xfit$Beta; Cov=xfit$Cov;
groups=dimnames(Beta)[[3]]
if(is.null(prior)){
prior=rep(1,length(groups));
names(prior)=groups;
prior=prior/sum(prior);
}
Sdet=VarMod[1]; Smax=VarMod[2];
# Returns the Posterior for each data-point in xpnts....
if(abs(sum(prior)-1)>1e-6) stop("Prior does not add to one")
# This works OK for now
# if(any(is.na(xpnts[,c("Pu.238","Pu.240","Pu.241","Pu.242")]))
# stop("NA not allowed in Data");
x=xpnts[,"Pu.240"]
y=xpnts[,c("Pu.238","Pu.241","Pu.242")];
lglik=array(0,dim=c(length(x),length(prior)));
colnames(lglik)=groups;
for(ii in groups){

```

```

Mu= Beta[,ii] %*% rbind(1,x,x^2);
Sp=sqrt(diag(Cov[,ii]));
St=pmin(sqrt(Sp^2 + exp(2*(Sdet-Mu))),Smax);
Cr= Cov[,ii]*outer(1/Sp,1/Sp)
logDCr=sum(log(eigen(Cr)$values));
sqlogdet= -0.5*logDCr - apply(log(St),2,sum);
y1=(t(y)-Mu)/St;
qterm = -0.5*apply(y1*solve(Cr,y1),2,sum);
lglik[,ii]= sqlogdet + qterm + log(prior[ii]);
}
post=exp(lglik);
post=post/apply(post,1,sum);
# if one is required to choose a class
if(classify){
  has.na=apply(is.na(post),1,any);
  zz=apply(-post,1,order)[1,];
  post[]=0; post[cbind(1:length(zz),zz)]=1;
  post[has.na,]=NA;
}
return(post)
}

#####
predfun=function(Xtrain,Xtest,prior=NULL,
  VarMod=c(Sdet=-Inf,Smax=5),linear=T,classify=F){
# Xtrain: dataset or fitted results that will be used to construct classifier
# Xtest: dataset that will be used to test classifier

if(all(names(Xtrain)[1:3]==c("Beta","Cov","regfits"))) xfit=Xtrain
else xfit=fitDbayes(Xtrain,linear=linear);
# This coding allows me to use an existing fit...
post=Dbayes(Xtest,xfit,prior=prior,VarMod=VarMod,classify=classify);
confmat=array(NA,dim=rep(dim(post)[2],2));
dimnames(confmat)=list(Actual=dimnames(post)[[2]],
Pred=dimnames(post)[[2]]);
for(i in 1:dim(post)[2]){
# Note: this returns "NaN" when there isn't anything in a category
confmat[,i]=tapply(post[,i],Xtest$reactor.series,mean,na.rm=T)
}
confmat[is.nan(confmat)]=NA;
return(confmat);
}
}

```

```

#####

#####
Crossval=function(data,pfun=predfun,cv.sets=lapply(1:dim(data)[1],max),
  prior=NULL,VarMod=c(Sdet=-Inf,Smax=5),linear=T,classify=F){
# predfun fits the model and then applies it to the cross validation sets
# data contains the data predfun uses. data will be split into a training
# set and test set
# val.sets: is a list with each component containing a set of indicies to
# be in a test (cross validation) set. So default is leave one out CV.

res0=pfun(data,data,prior=prior,VarMod=VarMod,linear=linear,classify=classify);
res1=array(0,dim=c(dim(res0),length(cv.sets)));
dimnames(res1)=list(rownames(res0),colnames(res0),names(cv.sets));
for(i in 1:length(cv.sets)){
  cat("Working on ",i,"\n");
  res1[,i]=pfun(data[-cv.sets[[i]],],data[cv.sets[[i]],],
  prior=prior,VarMod=VarMod,linear=linear,classify=classify);
}
av.cm=apply(res1,1:2,mean,na.rm=T);
sd.cm=apply(res1,1:2,sd,na.rm=T);
return(list(cv.sets=cv.sets,cv.runs=res1,Alldat.cm=res0,cv.cm=av.cm,cv.sd=sd.cm));
}
#####

#####
select.sets=function(data,ratio=.10,Nrep=100){
# dataset to use in cross validation exercise
# ratio Proportion of each group that is to be placed in the test data..
Rseed=.Random.seed;
y=data$reactor.series;
Ns=round(pmax(table(y)*ratio,1));
ygrouped=vector(length(levels(y)),mode="list");
names(ygrouped)=levels(y);
for(i in names(ygrouped)) ygrouped[[i]]=seq(y)[y==i];
cvset=vector(Nrep,mode="list");
for(j in 1:length(ygrouped)){
  for(i in 1:Nrep) cvset[[i]]=c(cvset[[i]],sample(ygrouped[[j]],Ns[j]));
}
attr(cvset,"random.seed")=Rseed;
return(cvset);
}
#####

```

```

#####
SelectByReactor=function(data){
# dataset to use in cross validation exercise
# ratio Proportion of each reactor that is to be placed in the test data..
y=data$reactor;
Ns=table(y);
ygrouped=vector(length(levels(y)),mode="list");
names(ygrouped)=levels(y);
for(i in names(ygrouped)) ygrouped[[i]]=seq(y)[y==i];
cvset=ygrouped;
return(cvset);
}
#####

#####
SimData=function(xfit,Pu.240=c(-5,0),VarMod=c(Sdet=-Inf,Smax=5),nmc=1000){
# this produces an "xdata" set of simulated values.....
Sdet=VarMod[1]; Smax=VarMod[2];
grps=dimnames(xfit$Beta)[[3]]; ng=length(grps);
np=length(Pu.240);
iseed=.Random.seed;
nobs=nmc*ng*length(Pu.240);
rseries=numeric(nobs);
Xtest=array(0,dim=c(nobs,4));
colnames(Xtest)=c("Pu.240","Pu.238","Pu.241","Pu.242");
icnt=0;
idx=cbind(rep(1:ng,np),rep(1:np,rep(ng,np)));
for(i in 1:(ng*np)){
jg=idx[i,1];
jp=idx[i,2];
kslc=icnt+1:nmc;
rseries[kslc]=jg;
Xtest[kslc,1]=Pu.240[jp];
mu = as.vector(xfit$Beta[,jg] %*% c(1,Pu.240[jp],Pu.240[jp]^2));
Sp=sqrt(diag(xfit$Cov[,jg]))
Crl= xfit$Cov[,jg]*outer(1/Sp,1/Sp);
St=pmin(sqrt(Sp^2 + exp(2*(Sdet-mu))),Smax);
Xtest[kslc,2:4]=mvrnorm(n=nmc,rep(0,length(mu)),Crl);
for(kk in 2:4) Xtest[kslc,kk]=mu[kk-1]+Xtest[kslc,kk]*St[kk-1];
icnt=icnt+nmc;
}
}
return(data.frame(

```

```

reactor.series=factor(rseries,levels=1:ng,labels=grps),
Xtest));
}

#####
IntDbayes=function(xfit,Pu.240=seq(from=-5,to=0,length=6),
  prior=NULL,VarMod=c(-Inf,5),nmc=1000,linear=T,classify=F,output=F){
# calculate confusion matrix for each Pu.240 value using MC Integration.
# xfit; Regression Model...
iseed=.Random.seed;
grps=dimnames(xfit$Beta)[[3]]; ng=length(grps);
confmat=array(NA,dim=c(ng,ng,length(Pu.240)));
dimnames(confmat)=list(Actual=grps,Pred=grps,Pu.240=round(Pu.240,2));
mcdata=vector(length(Pu.240),mode="list");
for(i in 1:length(Pu.240)){
xsim=SimData(xfit,Pu.240=Pu.240[i],VarMod=VarMod,nmc=nmc);
confmat[,i]=predfun(xfit,xsim,
  prior=prior,classify=classify,VarMod=VarMod,linear=linear);
if(output) mcdata[[i]]=xsim
}
res1=list(xfit=xfit,prior=prior,VarMod=VarMod,classify=classify,nmc=nmc,iseed=iseed);
res=list(Pu.240=Pu.240,confmat=confmat,mcdata=mcdata,Arguments=res1);
return(res);
}

#####
plot.IntDbayes=function(res){
Pu.240=exp(res$Pu.240);
error=res$confmat[1,];
nx=dim(error)[1];
for(i in 1:nx) error[i,]=1-res$confmat[i,i];
# error=cbind(t(error),total=apply(error,2,mean));
error=t(error);
matplot(Pu.240,error,type="l",log="x",xlab="Pu.240/Pu.239",ylab="Error",
  main="Classification Error",lty=1:nx,col=1:nx,lwd=3);
legend("topright",colnames(error),lty=1:nx,col=1:nx,lwd=3);
invisible();
}

```

References

- Betti M., T. G. (1999). Use of Secondary Ion Mass Spectrometry in Nuclear Forensics analysis for the Characterization of Plutonium and Highly Enriched Uranium Particles. *Analytical Chemistry Volume 71*, 2616-2622.
- Borg L., H. I. (2013). *Forensic Analysis of samples from the Nuclear Fuel Cycle, LLNL-TR-636458*. Livermore: Lawrence Livermore National Lab.
- Burr T., C. W. (2005). Assessing confidence in inferring reactor type and fuel burnup: A Markov chain Monte Carlo approach. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. Volume 555, Issues 1–2*, 426-434.
- Charlton W., e. a. (2000). Operator Declaration Verification Technique for Spent Fuel at Reprocessing Facilities. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 98.
- Charlton W., S. W. (2001). Monitors for the Prediction of Alternate Nuclear Material Concentrations for Pressurized Water Reactor Spent Fuel. *Nuclear Technology Vol 136*, 24-36.
- Chirayath S., H. G. (2009). Neutronic and nonproliferation characteristics of (PuO₂-UO₂) and (PuO₂-ThO₂) as fast reactor fuels. *Nucl. Eng. and Des.*, 239, 1916-1924.
- Chirayath S., J. O. (2015). Trace fission product ratios for nuclear forensics attribution of weapons-grade plutonium from fast and thermal reactors. *Sci. & Global Secur.*, 23, 48-67.
- Collins B., J. S. (2010). *Reactor Burnup Simulations and the Impact on Plutonium Forensic Signatures. PNNL-19454*. Richland, WA: Pacific Northwest National Lab.
- Dayman K., J. C. (2014). Characterization of Used Nuclear Fuel with Multivariate Analysis for Process Monitoring. *Nucl. Instrum. Methods Phys. Res., Sect. A*, 735, 624-632.
- DHS. (2022, 03 14). *Department of Homeland Security*. Retrieved from National Technical Nuclear Forensics Center: <https://www.dhs.gov/national-technical-nuclear-forensics-center>
- DOD. (2018). *Nuclear posture review report*. Washington, DC: Department of Defense.
- DOE. (2022, 03 14). *Department of Energy - National Nuclear Security Administration*. Retrieved from Counterterrorism and Counterproliferation: <https://www.energy.gov/nnsa/counterterrorism-and-counterproliferation>
- DOE. (January 1993). *DOE Fundamentals Handbook - Nuclear Physics and Reactor Theory Volume 2 of 2*. Washington, DC: US Department of Energy.
- Gauld I., L. K. (1995). *Verification and Validation of the ORIGEN-S Code and Nuclear Data Libraries*. tomic Energy of Canada Limited (AECL).

- Hellesen C., e. a. (2017). Nuclear spent fuel parameter determination using multivariate analysis of fission product gamma spectra. *Ann. Nucl. Energy*, 110, 886-895.
- Hoover A., e. a. (2009). *Measurement of plutonium in spent nuclear fuel by self-induced x-ray fluorescence*. Los Alamos: LANL - LA-UR-09-03481.
- IAEA. (2015). *Nuclear Forensics in Support of Investigations, Implementing guide*. Vienna: International Atomic Energy Agency.
- IAEA. (2017). *Identification of high confidence nuclear forensics signatures - IAEA TECDOC 1820*. Vienna: International Atomic Energy Agency.
- IAEA. (2018). *Development of A National Nuclear Forensics Library: A System For The Identification of Nuclear or Other Radioactive Material Out of Regulatory Control - IAEA-TDL-009*. Vienna: IAEA.
- IAEA. (2020). *Nuclear forensics: beyond the science, IAEA TECDOC 1896*. Vienna: International Atomic Energy Agency.
- Jardine L. (2005). *Radiochemical Assays of Irradiated VVER-440 Fuel for Use in Spent Fuel Burnup Credit Activities - UCRL-TR-212202*. Livermore: Lawrence Livermore National Laboratory.
- Jones A., P. T. (2014). Classification of spent reactor fuel for nuclear forensics. *Anal. Chem.*, 86, 5399-5404.
- Jordheim D. (1990). *CHAIN.238DJ: A Computer Code for Calculating Pu-238 Production, Quality, and Impurity Levels in the Np-237 Transmutation Chain*. Corvallis: Oregon State University.
- Joshi J., T. A. (2019). *Absolute Measurement of 137Cs and 134Cs/137Cs and 154Eu/137Cs Ratios to Verify University of Texas TRIGA Reactor Spent Fuel BU*. Los Alamos: Los Alamos National Lab - LA-UR-19-29132.
- Kitcher E., J. O. (2019). Sensitivity studies on a novel nuclear forensics methodology for source reactor-type discrimination of separated weapons grade plutonium. *Nucl. Eng. Tech.*, 51, 1355-1364.
- Lamarsh, J., & Baratta, A. (2001). *Introduction to Nuclear Engineering*. Saddle River: Prentice Hall.
- Luksic A., P. R. (2012). *FY11 Summary Report: Plutonium Signatures. PNNL-21995*. Richland, WA: Pacific Northwest National Lab.
- Makarova T., e. a. (2008). Destructive Analysis of the Nuclide Composition of Spent Fuel of WVER-440, WVER-1000 and RBMK-1000 Reactors, Khlopin Radium Institute. *Radiochemistry Vol 50, No 4*, 414-426.
- Mayer K., M. W. (2005). Nuclear forensics—a methodology providing clues on the origin of illicitly trafficked nuclear materials. *Analyst volume 130*, 433-441.
- Mayer M., W. &. (2000). Age determination of plutonium material in nuclear forensics by thermal ionisation mass spectrometry. *Fresenius' Journal of Analytical Chemistry volume 366*, 234-238.
- Michel-Sendis F., e. a. (2017). SFCOMPO-2.0: An OECD NEA Database of Spent Nuclear Fuel isotopic Assays, Reactor Design Specifications and Operating Data. *Annals of Nuclear Energy 110*, 779-788.
- Moody K., G. P. (2015). *Nuclear Forensic Analysis*. Boca Raton: CRC Press.

- National Security Council. (2021). *Interim National Security Strategic Guidance*. Washington: White House.
- NEA. (2011). *Spent Nuclear Fuel Assay Data for Isotopic Validation, State-of-the-Art Report*. Paris: NEA/OECD.
- Nicolaou G. (2017). Discrimination of spent nuclear fuels in nuclear forensics through isotopic fingerprinting. *Ann. Nucl. Energy*, 72, 130-133.
- Nikolaou G., S. B. (2019). Discrimination of weapons-grade plutonium from thermal reactors in nuclear forensics. *J. Radioanal. Nucl. Chem* 322, 513-518.
- Nuclear Power. (2022, 03 12). *Neutron Energy*. Retrieved from Nuclear Power: <https://www.nuclear-power.com/nuclear-power/reactor-physics/atomic-nuclear-physics/fundamental-particles/neutron/neutron-energy/>
- ORNL. (2011). *ORIGEN-ARP: Automatic Rapid Processing for Spent Fuel Depletion, Decay, and Source Term Analysis - ORNL/TM-2005/39 Version 6.1*. Oak Ridge: Oak Ridge National Laboratory.
- Osborn J., e. a. (2018). Attribution of Chemically Separated Plutonium. *Nuclear Technology*, 201:1, 1-10.
- Osborn J., e. a. (2018). Computational and experimental forensics characterization of weapons-grade plutonium produced in a thermal neutron environment. *Nuclear Engineering and Technology*, 50, 820-828.
- Parker D., M. R. (2001). *FISPIN10 Validation Review*. British Nuclear Fuels (BNFL).
- Radulescu G., I. C. (2010). *SCALE 5.1 Predictions of PWR Spent Nuclear Fuel Isotopic Compositions, ORNL/TM-2010/44*. Oak Ridge: Oak Ridge National Lab.
- Sternat M. (2012). *Development of Technical Nuclear Forensics for Spent Research Reactor Fuel*. College Station: Texas A&M University.
- Swinney M. (2015). *Experimental and Computational Assessment of Trace Nuclide Ratios in Weapons Grade Plutonium for Nuclear Forensics Analysis*. College Station: Texas A&M University.
- Tandon L., H. E. (2008). Nuclear, Chemical, and Physical Characterization of Nuclear Materials. *Journal of Radioanalytical and Nuclear Chemistry Vol 276*, 467-473.
- Toffer H., K. A. (1970). *Experimental Isotopic Analysis of Point Exposure Data in Hanford Production Reactor Fuels DUN-7243*. Douglas United Nuclear.
- Wallenius M., P. P. (2000). Origin Determination of Plutonium Material in Nuclear Forensics. *Journal of Radioanalytical and Nuclear Chemistry vol 246*, 317-321.
- Wolf S., B. D. (2000). *Analysis of Spent Fuel Nuclear Samples from Three Mile Island and the Quad Cities Reactors: Final Report*. Chicago: Argonne National Lab.
- Yamamoto T. (2009). *Compilation of Measurement and Analysis Results of Isotopic Inventories of Spent BWR Fuels*. Japan Nuclear Energy Safety Organization.