Ouachita Baptist University
Scholarly Commons @ Ouachita

Scholars Day Conference | Virtual Scholars Day 2020

May 1st, 12:00 AM

Analyzing Flash X-Ray Machine Diagnostics

Sebastian Bustillo
Ouachita Baptist University

Asaph Camillo
Ouachita Baptist University

Jake Morris
Ouachita Baptist University

Follow this and additional works at: https://scholarlycommons.obu.edu/scholars_day_conference

Part of the Computer Sciences Commons, and the Physics Commons

Bustillo, Sebastian; Camillo, Asaph; and Morris, Jake, "Analyzing Flash X-Ray Machine Diagnostics" (2020). *Scholars Day Conference*. 11.
https://scholarlycommons.obu.edu/scholars_day_conference/2020/natural_sciences/11

This Poster is brought to you for free and open access by the Carl Goodson Honors Program at Scholarly Commons @ Ouachita. It has been accepted for inclusion in Scholars Day Conference by an authorized administrator of Scholarly Commons @ Ouachita. For more information, please contact mortensona@obu.edu.

ANALYZING FLASH X-RAY MACHINE DIAGNOSTICS

SEBASTIAN BUSTILLO, ASAPH CAMILLO, JAKE MORRIS

Faculty Advisor: Jeffery Sykes, PhD

Abstract. The Cygnus Dual Beam Radiographic Facility consists of two flash x-ray machines, Cygnus 1 and 2. Seamless performance of this machines is critical to the maintenance of the United States stockpile of nuclear weapons. Since these SubCritical experiments cost about $104 million, when something malfunctions millions of dollars are lost. As a result, its imperative to asses the performance of the machine from diagnostics collected by the different sensors. Its performance it is measured by the level of radiation dose a shot obtains. Utilizing exploratory data analysis, interesting trends were found and a 74.1% level of correctness was achieved when predicting dosage.

1. Introduction

1.1. **Industry Partner.** This project was performed in conjunction with the Nevada National Security Site (NNSS). The NNSS works closely with the United States government and the Department of Energy. Their primary mission is to ensure the safety and defense of the United States and its citizens. This is achieved through 3 fields of focus. The first of these is nuclear weapons science. In this field, scientists study the stockpile of nuclear weapons that the United States has at their disposal. Most of these are a remainder of what was created in the Cold War era but never was utilized. The focus for the NNSS in this area is not to create more weapons and build upon the stockpile. Instead, a variety of tests and experiments are performed on the pre-existing weapons and materials that the weapons are composed of. They seek to better understand the weapons, the materials they are made of and how they have changed over time. Most of the weapons are from the Cold War era and therefore can be up to 30 or 40 years old. Understanding of the weapons is essential in ensuring that the weapons do not become unstable or dangerous, while also ensuring that they remain effective. The tests that are performed that analyze the weapons include complex diagnostic analysis, computational simulations and in-depth engineering analysis.

Figure 1. Location of test site

The second field of focus for the NNSS is related to global and homeland security programs. The organization is closely involved in efforts of nuclear nonproliferation around the world. Nuclear nonproliferation aims to reduce the amount of active nuclear weapons around the world. The goal of the NNSS in this area is to work with other nations in negotiations and meetings to work towards disarming as many nuclear weapons as possible. They encourage the study of nuclear materials to be directed towards nuclear energy production and not towards weapons that have an immense potential to cause mass destruction throughout the world. The NNSS specializes in nuclear and radiological emergency response efforts as a part of the global and homeland security programs. As a part of this program, the emergency response team goes to large events like the Super Bowl and the Air Force Academy graduation to perform tests and check for potential nuclear, chemical or biological threats. This team is the experts in the country for responding to these types of threats and are therefore the primary knowledge when training first responders. The last area of the NNSS focuses on environmental studies. Instead of studying materials or responding to emergencies, this team analyzes water, air and ground pollution levels throughout the nation. They also work towards the efforts of cleaning up radioactive waste and finding ways to permanently dispose of it.

1.2. **Cygnus Machine.** This project is a part of the stockpile stewardship area that focuses on studying pre-existing nuclear weapons and the materials they are composed of. The subject of ourThe machine consists of two x-ray sources, called Cygnus 1 and Cygnus 2. Our team focused primarily on Cygnus for this project. The Cygnus sources primarily consist of a Marx generator, inductive voltage adder (IVA), and a rod pinch diode. The IVA supplies the necessary power to the Marx generator which in turn creates a high energy pulse that is fed through the rod pinch diode. The diode then converts that energy into high energy x-rays that leave the cygnus machine. Once fired the x-rays enter the containment vessel where a sample of plutonium is located. The plutonium can either be static or dynamic. At times, the material is simply placed in the vessel, and other times the material is detonated to analyze how it acts and moves during the explosion. The x-rays pass through the containment vessel where they either hit material and are absorbed, are deflected, or pass through the vessel untouched. Those rays that leave the containment vessel then pass through a scintillator that produces a visible light when struck with a high energy ray. The visible light travels into the camera systems at the end of the apparatus which produces an image. This image is what can be studied by the scientists at NNTS. It provides analysis of the material at a depth instead of just surface analysis. project is a high-energy x-ray radiography machine called Cygnus, that the NNSS uses to analyze plutonium.

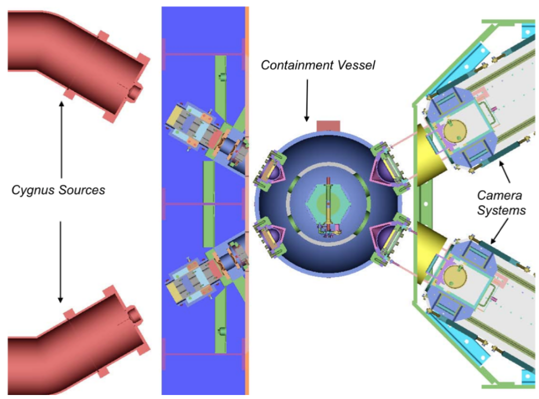

Figure 2. Diagram of Cygnus machine

The machine consists of two x-ray sources, called Cygnus 1 and Cygnus 2. Our team focused primarily on Cygnus for this project. The Cygnus sources primarily consist of a Marx generator, inductive voltage adder (IVA), and a rod pinch diode. The IVA supplies the necessary power to the Marx generator which in turn creates a high energy pulse that is fed through the rod pinch diode. The diode then converts that energy into high energy x-rays that leave the cygnus machine. Once fired the x-rays enter the containment vessel where a sample of plutonium is located. The plutonium can either be static or dynamic. At times, the material is simply placed in the vessel, and other times the material is detonated to analyze how it acts and moves during the explosion. The x-rays pass through the containment vessel where they either hit material and are absorbed, are deflected, or pass through the vessel untouched. Those rays that leave the containment vessel then pass through a scintillator that produces a visible light when struck with a high energy ray. The visible light travels into the camera systems at the end of the apparatus which produces an image. This image is what can be studied by the scientists at NNTS. It provides analysis of the material at a depth instead of just surface analysis.

1.3. **Problem and Goal.** The problem that we are working to solve on this project revolves around the Cygnus machine failing. Cygnus is comprised of many components that all have to operate correctly, and in the correct sequence in order for the machine to perform correctly. When operating successfully, the machine will produce a radiation dosage of about 4 radians. When unsuccessful, an inadequate amount of radiation is produced which leads to an image that is not as clear and easy to read and analyze. It is important for the Cygnus machines to perform correctly in both the small and large scale experiments. The small scale experiments happen twice a day. This is due to the extensive time and energy it takes to perform an experiment. When fired, the diode electrodes found in the rod pinch diode are destroyed. The diode must be removed, cleaned and a new electrode inserted before another experiment can be performed. The primary purpose of work on this project is in regards to the large scale experiments. It is imperative that Cygnus does not fail when it is time to perform a large experiment. These experiments only happen once or twice every year, can cost tens of millions of dollars, and cannot simply be reset or redone if a mistake happens due to detonation. Therefore, the primary goal is to be able to predict when a bad shot might occur so that it can be prevented. This could potentially save the NNSS millions of dollars and months of time. This is done through analysis of large amounts of data, provided by the NNSS, from diagnostic tools connected to Cygnus. As this project progressed it became clear that it was necessary to alter the goal. Creating a predictive model based on the data is of a scope that is beyond the semester of work that we were able to perform. Instead, our goal is to provide the team at NNSS with a quicker way of identifying a bad shot. It can take hours to read the radiation dose that is produced by the machine, but if a relationship between the data and shot success can be found, a bad shot can be instantaneously identified by the diagnostic tools.

2. Data

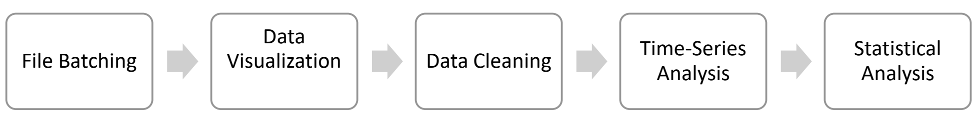

Figure 3. Steps taken in solving the problems

2.1. **File Batch Processing.** The raw data was received in a format that was hard to manipulate. In order to, organize the data on a more meaningful way and automate the whole process it was crucial to have all the shots arranged in sequential order. Some of this batch processing was done using Micorsoft's PowerShell, this command-line interface allows the modification, duplication, mobilization, and creation of files all at once. PowerShell proved to be useful in taking care of the repetitive, uninteresting job. Then the data was all sorted out by shot number, each shot number contained all its corresponding diagnostics.

After some time working with the structured data, it was discovered by just looking at different diagnostics of the same shot, that it was pointless to have it this way because they were all distinct (Fig.2). Some varying in width, range, or even the intensity of the shot. It was necessary to go back and reorganize the data again.

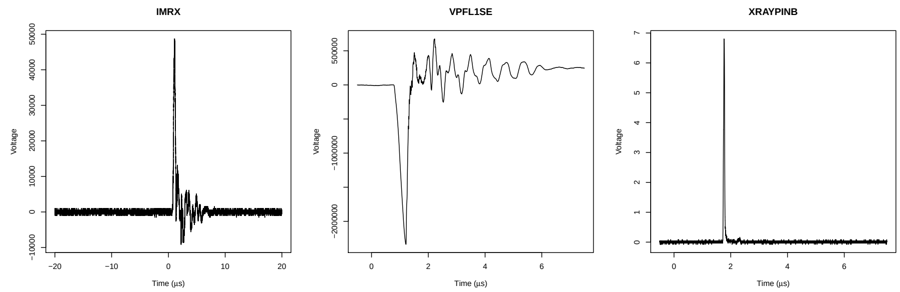

Figure 4. Diagnostics of shot 2317

Now the data was rearranged into a different structure. All diagnostics were put under one single directory. Then each diagnostic sub-directory had all the shots corresponding to that diagnostic specifically. It was now possible to compare different shots and their doses.

2.2. **Data Visualization.** One of the techniques that data scientists rely a lot on is data visualization, since is the first step towards data analysis. As the name itself implies, the data is visualized in order to discover any anomalies on the signal.

MATLAB and R were utilized for this step. By comparing their performance and the different variety of libraries available, it seemed like R was going to be used for the remaining of the project. Mainly, because of all its statistical learning, data visualization, cleaning, and smoothing libraries. On top of that, R is an open-source programming language, there is no cost associated with its use.

An R script was developed to perform the visualization of voltage and current traces for the different shots. Then these were saved as PDF files, which were used in the exploratory data analysis step (EDA). Over 600 PDF files were created in order to analyze them. After hours of looking at tons of plots and noticing the variances between them. A substantial difference was noticed on the range of a few shots on the diagnostic XRAYPINB, the doses were compared. Unsurprisingly, a weak correlation seemed to appear between the range of the trace and the radiation dose.

3. Methodology

It was a little bit challenging to find the right approach to the problem. However, after some analysis of the potential ways to tackle this problem it was determined that the approach was going to have other integral parts to it. The steps taken towards finding a solution to the problem are highlighted in figure 3.

3.1. **Data Cleaning.** As mentioned earlier, a weak but resounding correlation was identified in the visualization process. Some doubt was raised to whether or nor it was possible to somewhat improve this correlation. With this in mind, it was determined that cleaning the data was the next step to follow in the quest of mapping some characteristic of the signal to the dosage.

Various different smoothing algorithms were tested, after some trials it was found that the following four smoothing techniques seemed to improve the resolution of the signal: **1) Moving Average Filter**: simple to understand, used with time series data to smooth out short-term noise/fluctuations **2) Cubic Spline**: used for curve fitting, less computational expensive, reduces data points **3) Gaussian Noise Elimination**: used in image processing, combines signal with a Gaussian function **4) Savitzky-Golay Filter**: digital filter, enhances/improves precision of the data without distorting the signal; mathematical complex, relatively computationally expensive

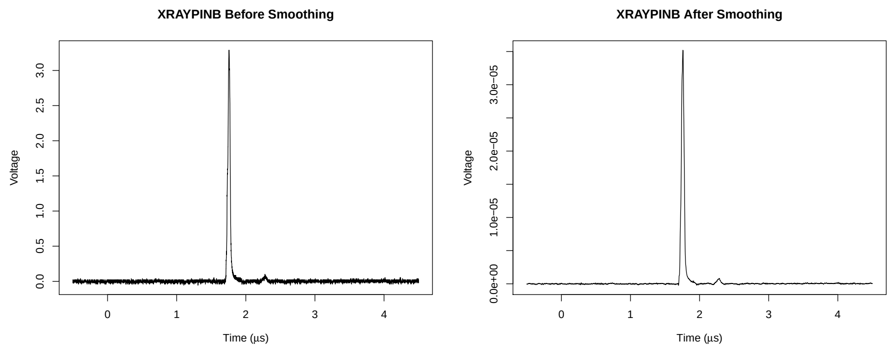

Figure 4. Signal after using Savitzky-Golay filter

The best resolution without a huge distortion of the data was provided by the Savitzky Golay with a filter length of 51 and an order of 4. Although this step becomes helpful whenever the signal is analyzed is not indispensable.

Some charactersitics were being observed in the data. For instance, all of diagnostics contained a peak, this was either negative or positive. Moreover, each signal seemed to be bounded to a baseline for most of the time, until some voltage or current would be turned on. This baseline value was calculated by using the Root Mean Square (RMS) of the interval having the most noise/fluctuations, the RMS is defined in equation 1.

$$RMS = \sqrt{\frac{1}{N}\sum_{1}^{N} y_i^2}$$

Equation 1. Root Mean Square

Further trying to simplify the problem as much as possible. A process of extracting the peak of each signal was implemented, which was denominated as Peak Cropping. Basically, the baseline was utilized as a threshold value. Anything below it was equaled to zero and since nothing happens until the signal passes the threshold value, then the signal was cropped between the two closest zeros to the signal, on the right and left.

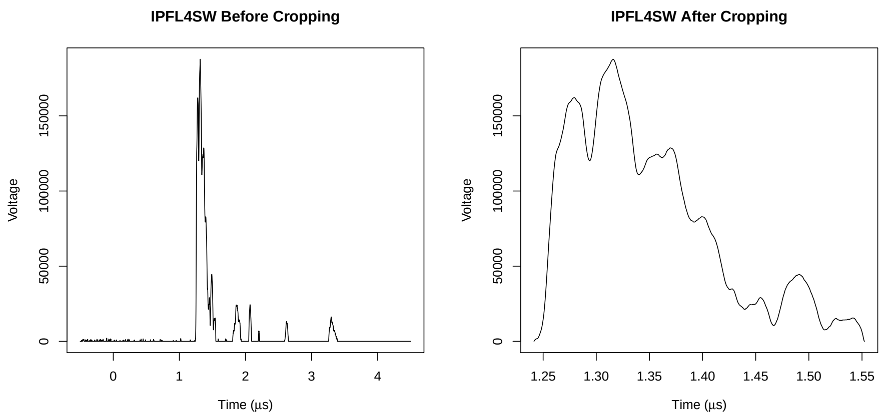

Figure 5. Demonstration of peak cropping

As shown in Fig.5, cropping the peak removes any extra data from the signal. Some concerns may arise regarding loosing valuable information from the trace. However, some in-depth analysis it was figured out that this process does give positive results.

Now everything is set to perform some statistical analysis to the signal. The following quantities were calculated for the signal: **1) Range**, maximum minus the minimum value in the signal **2) RMS**, root mean square value of the signal **3) MIN**, minimum value of the signal **4) Mean**, simple arithmetic mean of the signal **5) SD**, standard deviation of the signal **6) AOS**, area under the signal **7) AAOS**, absolute area under the signal

The time series of the different statistical quantities were plotted. Some of them seemed to show no trends whatsoever. Nonetheless, the range time series plot showed a very intriguing trend in the signal.

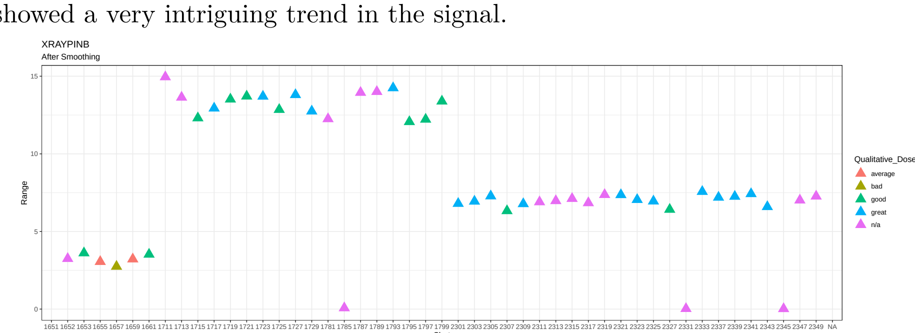

Figure 6. Range time series

Figure 6. shows the trend found in range variable for the diagnostic XRAYPINB. A clustering seemed to appear in the three different set of shots analyzed. Also, notice that each data point is color coded. The shot's quality was divided into one of four different categories as shown in Table 1.

Great Shot	dose$\geq$ 4.5
Good Shot	$4.5 > \text{dose} \geq 4.0$
Average Shot	$4.0 > \text{dose} \leq 3.5$
Bad Shot	dose < 3.5

Table 1. classification of shots

After noticing that each cluster of shots appeared to have some sort of relationship, dosage was plotted as a function of each statistic calculated for the signal. The results for some of the diagnostics were very satisfying, since what was found in the EDA step was shown to be accurate (Figure 7).

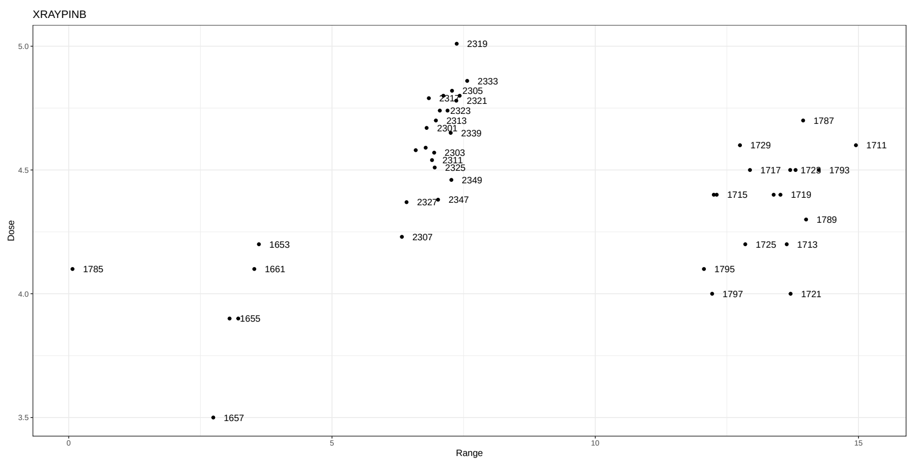

Figure 7. Dose as a function of range

Once more, the trend is visible. Notice that that a very linear fit could be utilized for the 2300s shots. On the other hand, the 1700s seem to be all over the place. This raises the question, does this deviation from the linear fit is a predictor will go bad? Very interesting question that is left for future investigation.

Lastly, finding the percentange of correlation was a way to be certain that this range quantity could become a potential predictor of the dose. Indeed, this proved to be true as Figure 8. shows the percent correlation between the different simple statistics calculated.

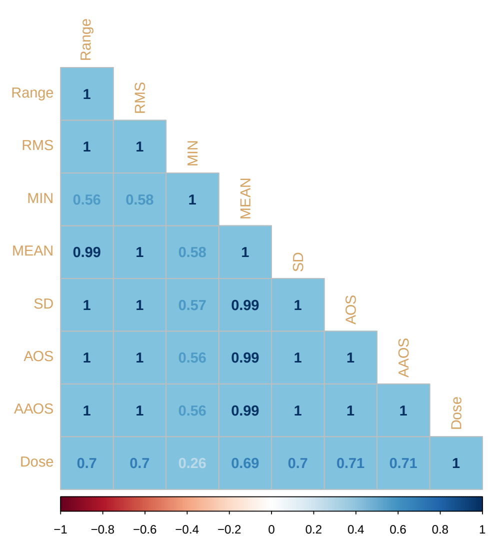

Figure 8. Correlation chart

Range as well as AAOS, AOS, and MAX seem to have a decent correlation to the dosage. More analysis needs to be done to compare different diagnostics.

4. Limitations

Throughout the project we faced a few obstacles that delayed the progress of the study. In the first couple weeks of work, our primary goal was downloading, organizing and visualizing the data. At first we utilized an online software called CoCalc. CoCalc is an online workspace that allows for the import of data, LaTex functionality, and being able to use programming languages like R to work with data. This software was relatively unknown to us, and so after a couple of days we decided to use MatLab for our data analysis because of familiarity. This choice improved the accuracy of several computations and visualization techniques as well as allowing us to do these things quickly. However, we eventually reached a point where some statistical analysis tools were intended to be used. These techniques we desired to use were not compatible with MatLab, so we decided to switch back to CoCalc and utilize the R programming language for our analysis. MatLab and R are not quite the same, but are similar enough in concept to be able to recreate past work easily.

Another problem that set us back a couple days work revolved around extraneous data. There were several shots, shot 1785 being one that looked drastically different than every other shot of that same diagnostic.

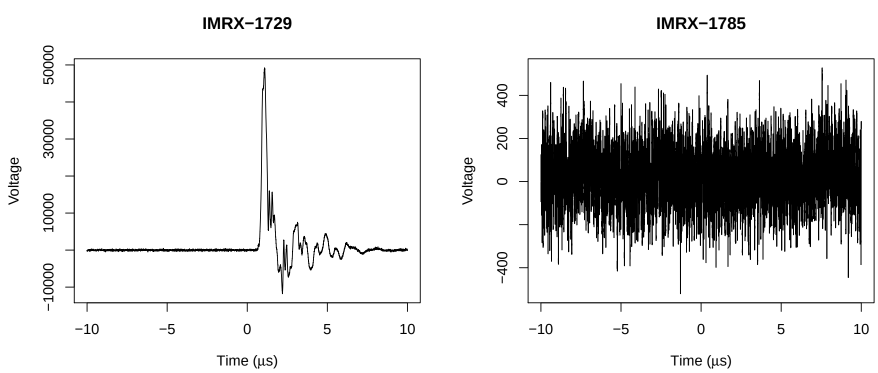

Figure 5. Normal and unusual signal of IMRX diagnostic

Our group spent almost a week analyzing these plots and what contributes to the drastic difference. We could not decide if this was relevant information or an outlier of some kind that could be discarded. In the end, it was concluded, based on the naming of the file and similar trends elsewhere in the data, that this plot was a result of a type of leakage Cygnus recorded. It was simply a result of a diagnostic tool on Cygnus 1 being on and recording when Cygnus 2 was executing. Therefore it was reading a machine that was not actively being used. Therefore it was decided this data was irrelevant and not useful, but took away several days of work that could have been spent elsewhere.

The final limitation of our project was with the data. A trend was discovered in the data of several diagnostics that was promising. We found linear trends between the range of a plot and the radiation dose of that particular shot, which could be extremely useful. However, we also discovered several different "clusters" of data. When looking at a time-series plot of the diagnostic where range is plotted versus dose, three separate groups emerged that all had linear trends within. The next step in following this idea would be to analyze more data in the same way while looking for this trend we have found. However, when we began needing this data, new data could not be provided due to COVID-19. Our industrial liaison did not have access to the site where the data was located and so we were unable to analyze more data.

5. Conclusions

Overall, the project yielded very promising results. First, it was determined that smoothing the data helped to improve the resolution of the signal. Secondly, noise from Cygnus 2 appeared to be leaking to some diagnostics of Cygnus 1 and vice versa. Thirdly, a clustering was identified within the data set that was given. Lastly, the dose of 14 shots was predicted with a 74.1% correctness. This is very encouraging to continue further analyzing the data. A simple linear model may be able to predict when a bad shot is going to happen, thus saving millions of dollars to the DOE.

6. Future Work

Some ideas to look at whenever more data becomes available is train a machine learning model to predict the dosage of the shot. Given the nature of the clusters, perhaps K-Nearest Neighbors could prove efficient to perform this. Moreover, a Komorov-Smirnov test could potentially help in determining any major differences between signals. This would help find other characteristic of the trace. A way to compare different diagnostic is neccesary, the observations cannot be based in one diagnostic solely.

7. Acknowledgments

PIC Math is a program of the Mathematical Association of America (MAA). Support for this MAA program is provided by the National Science Foundation (NSF grant DMS-1722275) and the National Security Agency (NSA). We also acknowledge the Nevada National Security Site (NNSS) and Dr. Marylessa Howard for all the support and willingness to work with us.

The Math & Computer Science Department at Ouachita Baptist University and Dr. Sykes for all the guidance.

References

[1] U.S. Department of Energy, Office of Inspector General. (2015) *Subcritical Experiment Activities at the Nevada National Security Site*
`https://www.energy.gov/sites/prod/files/2015/07/f24/OAS-L-15-08.pdf`

[2] Corrow, G., et al. *"Cygnus Performance in Subcritical Experiments." Cygnus Performance in Subcritical Experiments (Conference) | OSTI.GOV, 1 Feb. 2008,*
`www.osti.gov/servlets/purl/926630`

[3] Press, William H. *"Savitzky-Golay Smoothing Filters" Computers in Physics, vol.12, no.6, 1998.*
`https://aip.scitation.org/doi/pdf/10.1063/1.4822961`