Nuclear Engineering PhD Preliminary Exam

Automated Variable Selection of Gamma-Ray Spectra by Utilization of LASSO and Elastic Net Techniques for Use in Nuclear Security Applications

North Carolina State University

Raleigh, North Carolina

February 21, 2019

Vincent A. DiNova

Presented to the Committee:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | |  | | |  | | --- | |  | | |  | | --- | |  | |
| Dr. Dimitry Anistrotov | Dr. Ge Yang | Dr. Steven Shannon |

|  |  |  |  |
| --- | --- | --- | --- |
|  | |  | | --- | |  | |  |

Dr. Ralph Smith

With Committee chair:

|  |  |  |  |
| --- | --- | --- | --- |
|  | |  | | --- | |  | |  |

Dr. Robin P. Gardner

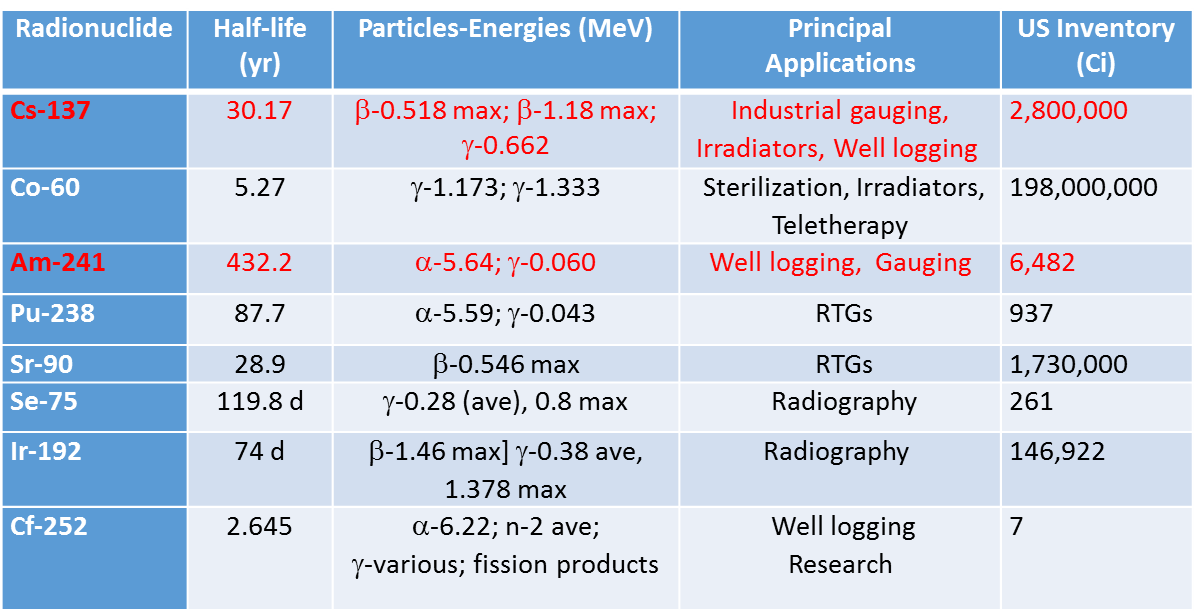
**CHAPTER 1**

**Introduction**

1.1 Background

Shortly after the tragedies of 9/11/2001, the National Academies of Science commissioned a study on the dangers of long-lived radioisotope sources.  The study concluded that there exist several commonly used sources that could potentially be used as a dirty bomb or terror weapon (Table 1-1).  The Consortium for Nonproliferation Enabling Capabilities (CNEC) was funded in 2014 to research innovative ways to address nuclear security problems including finding suitable replacements for dangerous radiological sources.

Table 1-1: NAS findings



Devices used in the oil well logging industry were identified as a major point of interest as they utilize high activity Cs-137 and AmBe sources for density, porosity, and elemental composition measurements.  A testing facility and benchmarking tool were designed and built at Kansas State University to test the viability of replacing traditional active sources with a D-T Pulsed Neutron Generator (PNG).  A PNG operates by receiving a signal to initialize a pulse firing sequence that propels Deuterons and Tritons on a collision path releasing 14.1 MeV neutrons. These high-energy neutrons are used as an alternative to AmBe neutrons in a traditional prompt gamma neutron activation analysis (PGNAA) application. PGNAA is a nondestructive method that relies on (n, γ) neutron capture, and (n, n’γ) neutron inelastic scattering reactions to produce gamma photons, each having distinct characteristics of the target nuclei. Using a near and far NaI scintillator detector, each spectral detector response can be analyzed for elemental composition.

PGNAA suffers from a low signal to noise ratio caused by the delayed activation of nuclei or neutron activation analysis (NAA). Neutron activation analysis utilizes the delayed gamma rays from radioactive daughters, while PGNAA exploits the prompt gamma rays (Fig. 1-1). The neutron cross sections for both prompt and delayed reactions compete and create a mixed signal that is often difficult to process. Additional sources of interference include:

1. γ rays produced by the PNG
2. γ rays from activation inside the detector medium (Gardner, 2000)
3. γ rays from background sources
4. γ rays produced by the activation of construction materials in the benchmarking tool.

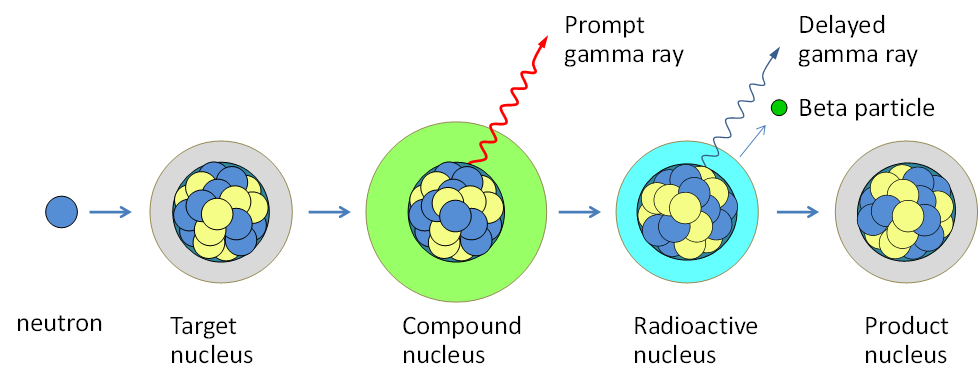


Figure 1-1 Prompt and delayed gamma ray emission process

The PNG offers a unique solution to this problem by exploiting the pulsing time responses with a digitizer allowing the prompt and delayed responses to be extracted and separated. This critical step allows for supervised machine learning variable selection techniques such as LASSO and Elastic Net to be applied to the prompt and delayed responses, offering on line analysis in a changing environment with improved capabilities over traditional linear least squares methods.

1.2 Benchmarking tool and facility

The benchmarking tool and design facility were designed and constructed at Kansas State University. The tool consists of near and far gamma and neutron detectors separated from a D-T PNG source by a 2” lead divider (Fig. 1-2). A CAEN 5730 digitizer acts to both send the pulse firing sequence to the PNG and to collect the responses from the gamma and neutron detectors (Fig. 1-3).

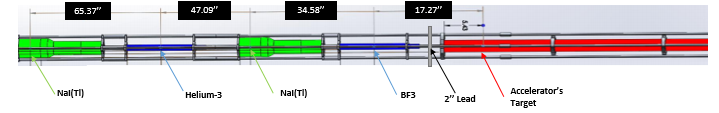


Figure 1-2: KSU benchmarking tool

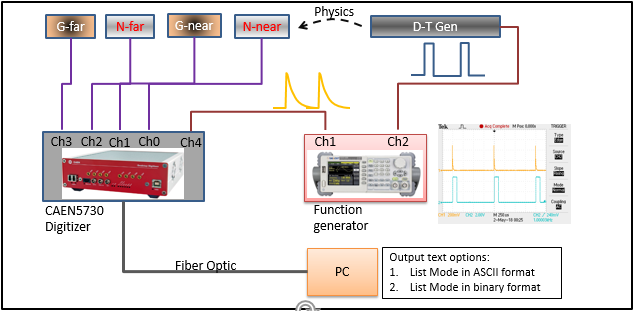


Figure 1-3: Data acquisition scheme

The design facility (Fig. 1-4) is located at King Hall Annex at Kansas State University. Special accelerator enabling systems were required to gain approval from the Kansas Safety Board to prevent inadvertent neutron production and entries into the facility. These safeguards include audio and video surveillance, controlled access points, intercommunication systems, warning lights, and detailed operating procedures to minimize unnecessary dosage to bystanders.

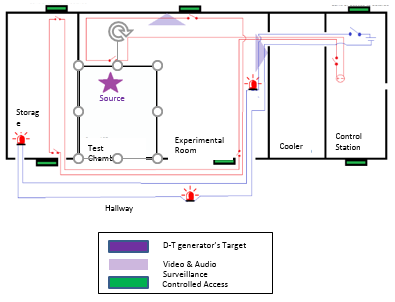


Figure 1-4: KSU design facility

The test chamber (Figs. 1-5, 1-6) is 6’8” high by 6’6” wide and 8’ deep with a total volume of roughly 2,500 gallons. These dimensions ensure that when fully filled with water, the test chamber presents an effectively infinite medium to the 14.1 MeV neutrons. During the data collection process, the benchmarking tool is loaded into the borehole tube and enclosed by a cap. Borated polyethylene (green material) has been fitted both inside and outside the test chamber to reduce neutron escape and dosage.

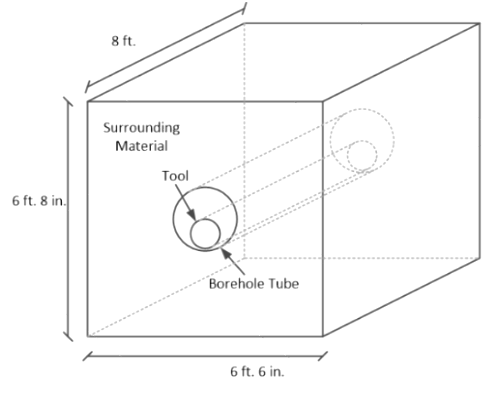


Figure 1-5: Test chamber

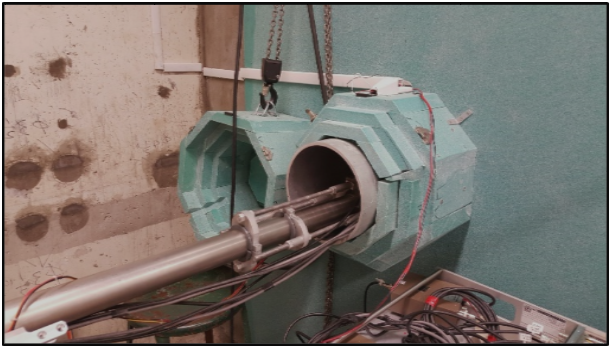


Figure 1-6: a. Open borehole tube b. Capped borehole tube

1.3 Monte Carlo Library Least Squares

In order to quantitatively analyze measured PGNAA γ spectra, monte carlo library least squares (MCLLS) has been shown to be an effective method over single peak analysis (Gardner, 1997). MCLLS requires extremely accurate forward model simulations to represent the expected pulse-height spectrum obtained with a PGNAA system with a known geometry and compositional makeup. Previous studies have demonstrated the effectiveness on bulk coal (Shyu et al., 1988, 1998) and PGNAA applications (Han, 2005 and Hou, 2017). The MCLLS approach consists of:

1. Generating pulse-height spectra with MCNP or similar coding packages using assumed or known geometry and compositions.
2. Utilizing each prompt γ-ray pulse-height spectrum as a library input variable for model selection.
3. Adjusting all non-linear parameters between simulated and experimental response to treat the problem as a sum of linear responses.
4. Perform a linear library least-squares (LLS) analysis.

Assuming all non-linear physical components are correctly adjusted, the library least-squares method treats an unknown sample as the sum of the products of an elemental amount with the library spectrum of each element for each channel as given by equation 1.1.

|  |  |  |
| --- | --- | --- |
|  |  | (1.1) |

Where,

* + is the counts per channel of an unknown spectrum
  + are the library spectra, or counts in channel of element
  + is random error in counts in channel

Equation 1.1 is solved for by minimizing the reduced Chi-Square given by equation 1.2.

|  |  |  |
| --- | --- | --- |
|  |  | (1.2) |
|  |  |  |

Where,

* + is the number of degrees of freedom
  + is random error in counts in channel
  + is the variance of the random error in counts in each channel *i*

1.4 Radioisotope Identification Devices

During this investigation on using LASSO and Elastic Net variable selection techniques on oil well logging devices, radioisotope identification devices (RIIDs) were targeted as another application that could be improved by the use of these algorithms. RIIDs are handheld instruments that are designed to identify radioactive materials by their emitted gamma ray energies. RIID algorithm development is necessary to ensure that potential threats are correctly identified by field agents and first responders. A study from 2003 concluded that only 30%-50% of medical, industrial, NORM, and threat isotopes were correctly identified (Blackadar et. al).

RIID algorithms measure similarities between measured characteristics of test spectrum and reference spectra, whether experimental or simulated. Many challenges emerge as a result of variations that occur naturally (Burr and Hamada, 2009) including:

* Absorbers can attenuate counts differentially with respect to energy, causing distorted pulse height distributions.
* Voltage calibrations can be nonlinear, and electronic gain shift effects can be nonlinear and unstable over time due to environmental effects
* Procedural protocols can complicate the required measurements. Temperature and spatial factors can impact the calibration and collection time necessary to collect accurate measurements.
* There can be hundreds of isotopes in the isotope library, each with numerous example spectra.
* Detector dead time and pulse pile up can distort the test spectra.

Due to inadequate RIID algorithm performance, the U.S. Department of Energy (DOE) requires trained spectroscopists to be on call to resolve alarms as rapidly as possible. It was identified that human experts outperform even the best RIID algorithms, leading to the study of supervised machine learning techniques established by experts to help address these shortcomings.

**CHAPTER 2**

**Nuclear Reactions**

Prompt gamma-ray neutron activation analysis (PGNAA) is a nondestructive, near real time technique used for bulk material identifications. PGNAA relies on neutron inelastic scatter and capture reactions to produce characteristic γ-rays used to identify minute amounts of elements in a bulk sample. Due to low cross sections for these reactions, background sources from natural radiation, activation of the NaI detector, and γ-rays from the decay of the neutron source a low signal to noise ratio (SNR) is common.

**2.1 Neutron Transport**

**2.1.1 Neutron Inelastic Scatter (n, n’γ)**

Neutron inelastic scatter involves an incoming neutron colliding with a target nucleus and exiting with less energy and at a different angle than it entered. The energy deposited on the target nucleus causes it to reach an excited state and rapidly releases a γ-ray to return to its normal energy state represented in equation 2.1 as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (2.1) |

The inelastic scattering reaction requires an incoming neutron to have enough energy to break the threshold energy derived from the Q value formula shown in equation 2.2.

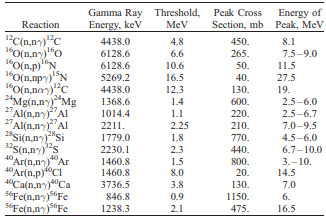
|  |  |  |
| --- | --- | --- |
|  |  | (2.2) |

Where,

* + is the kinetic energy of the incident neutron
  + is the energy of the target nucleus’ first excited level
  + is the atomic number of the target nucleus

The PNG is advantageous for this decay scheme, as the 14.1 MeV neutrons allow a greater number of isotopes to undergo inelastic scatter and unlock higher energy states. Table 2-1 (Kim et, al. 2006) below lists some common inelastic scatter threshold energies and cross sections.

Table 2-1: Non-elastic scattering reactions and threshold energies



**2.1.2 Neutron Capture (n, γ)**

Neutron capture, also denoted as (n, γ), can occur over a wide range of energies and has the highest probability at thermal energies. The (n, γ) reaction begins when a neutron interacts with a target nucleus and is absorbed. The newly formed nucleus is placed in an excited state, and in order to form a new ground state, at least one γ photon is emitted as shown in eq. 2.3 below.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  | (2.3) |

Each nucleus (apart from Helium-4) gives off a distinct signature of intensities and energies, allowing for the identification of the sample from the γ photon emissions.

**2.2 Photon Transport**

**2.2.1 Photon Reactions**

The benchmarking tool analysis relies on three main photon reactions: photoelectric absorption, Compton scattering, and pair production. Although there are other photon reactions, none are an important focus to this work.

**Photoelectric Absorption**

During photoelectric absorption, a photon interacts with a target atom’s electron, departs all of its energy, disappears, and ejects the electron from its bound shell. The photoelectron carries an energy given by equation 2.4 and illustrated by figure 2-1.

|  |  |  |
| --- | --- | --- |
|  |  | (2.3) |

Where,

* + is the binding energy of the photoelectron in its original shell
  + is the energy of the exited photoelectron
  + is the energy of the incoming photon

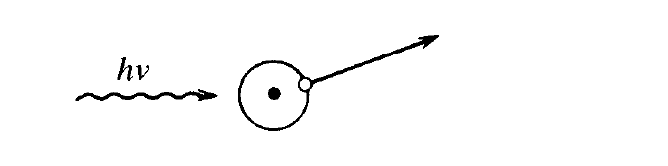
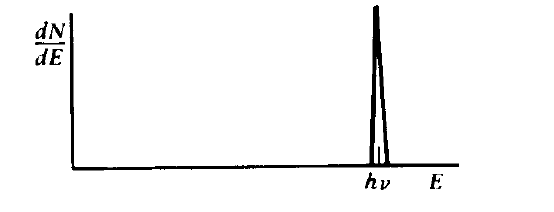


Figure 2‑1: Depiction of the Photoelectric Effect

**Compton Scattering**

Compton scattering takes place when a photon interacts with an electron in an atom, it deflected while imparting some of its original energy, and ejects an electron from its orbit. Depending on the scatting angle, the energy transferred to the electron can range from zero to a large fraction of the total photon energy. The energy of the scattered photon and kinetic energy of the scattered electron can be calculated if the energy of the incident photon and incident angle are known by:

|  |  |  |
| --- | --- | --- |
|  |  | (2.4) |

Where,

* + is the scattered photon energy
  + is the energy of the incident photon
  + is the scattering angle in the lab frame
  + is the mass of the electron
  + is the speed of light

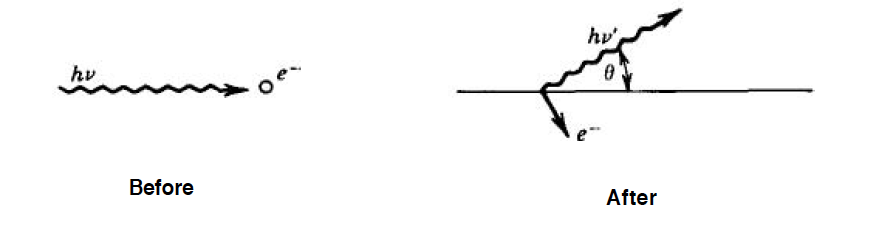


Figure 2-2: Depiction of the Compton Scattering Process

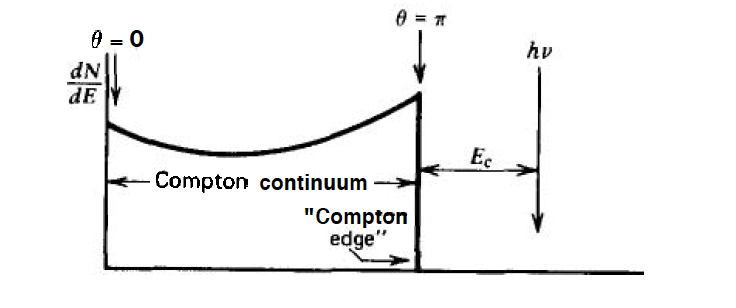


Figure 2-3: Illustration of the Compton Edge and Continuum

**Pair Production**

When an incoming photon exceeds 1.02 MeV, a photon can disappear and create an electron positron pair. All additional energy carried by the photon above 1.02 MeV is converted into kinetic energy shared by the positron and electron. Although possible at any energy above 1.02 MeV, pair production does not become the dominant reaction until photon energies exceed 5 MeV as shown in Figure 2-5.

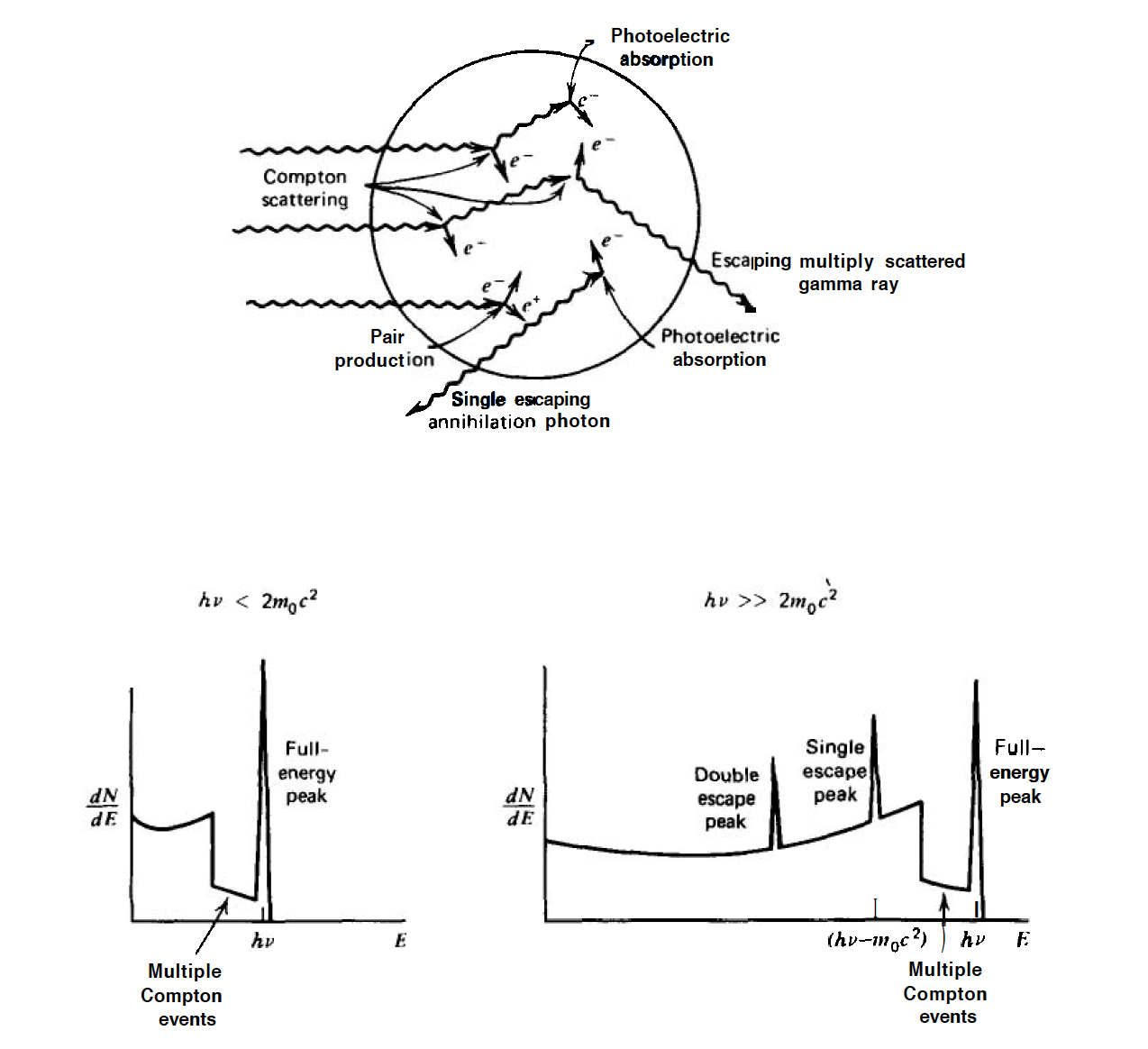


Figure 2-4: Depiction of interaction of gamma-rays with a detector medium and their resultant spectra

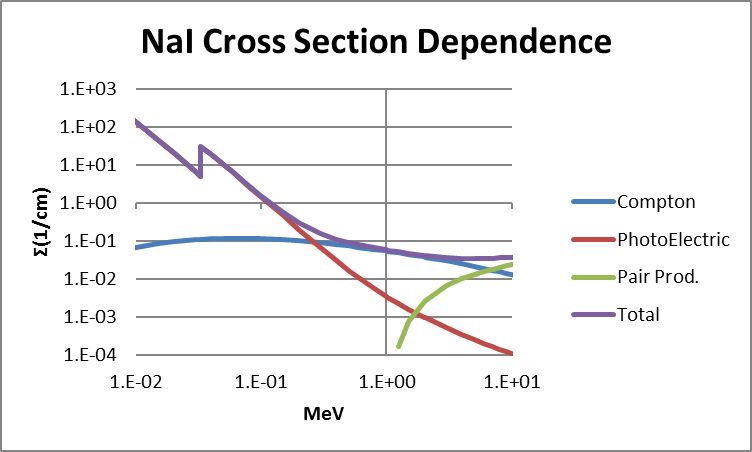


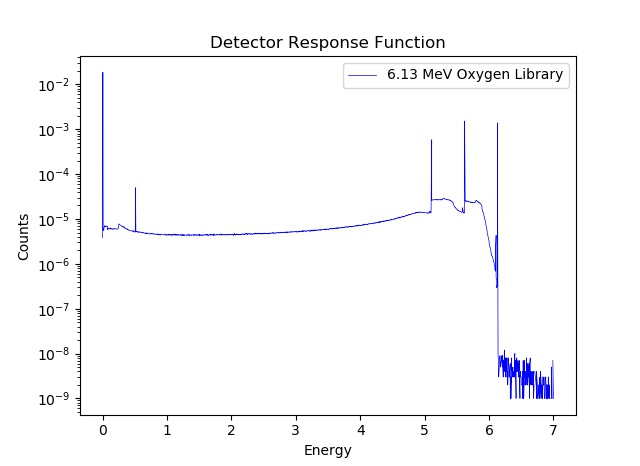
Figure 2-5: Energy dependence of photon interactions in NaI

**2.3 Detector Response**

Incident radiation interacting with the detector are modeled in MCNP 6.1 and carry distinct physical signatures (Knoll, 2011). The first order physics alone do not provide a genuine representation of the response from a detector. In order to accurately reproduce a detector response, a detector response function (DRF) must be applied. A detector response function is the expected detector response provided an incoming particle (Gardner and Sood, 2004). During this stage, all non-linear characteristics are removed, allowing the problem to be dealt with by linear methods.

**2.3.1 Spectral Features**

Figure 2-2 will serve as an example of the expected features produced by a monoenergetic photon on a detector. The response simulates a 0% resolution detector response from an incoming 6.13 MeV Oxygen-16 decay γ-ray.



F

B

E

H

C

G

D

A

Figure 2-2: Infinite resolution detector response

**A: Full Energy Peak**

The full energy peak is equal to the incoming photon energy, 6.13 MeV for this example. The pulse signals that produce this signature occur when the particle enters the detector and deposits all of its energy with no escaping secondary particles. This can occur in a singular photoelectric absorption or by a series of reactions. For this reason, the full energy peak can be referred to as the “photoelectric peak”.

**B: Compton Edge**

Not all events result in full energy deposition. The Compton edge occurs once a photon undergoes a Compton scattering reaction, and then exits the detector without depositing additional energy. The scattering angle determines the energy that is deposited in the detector, ranging from 0° to 180°. From equation 2.4, the maximum energy deposited by a Compton scattering identified as the Compton edge is

|  |  |  |
| --- | --- | --- |
|  |  | (2.5) |

**C: Single Escape Peak**

In the event a photon creates a pair production reaction, an electron-positron pair is created. The positron will then annihilate inside the detector, creating a 0.511 MeV photon. If the annihilation photon exits the detector without depositing its energy and the full energy of the incoming photon is deposited inside the detector, the result is the single escape peak. The energy of the peak is equal to the full energy peak minus the resting mass energy of an electron, 0.511 MeV.

**D: Double Escape Peak**

The double escape peak is the result of both annihilation photons escape the detector. In this way, the resulting energy is equal to that of the full energy peak minus the resting mass energy of two electrons, 1.022 MeV.

**E: Annihilation Peak**

In the event that the incoming particle interacts with a material outside of the detector by a pair production reaction, a resulting annihilation photon can enter the detector and deposit its energy. The annihilation peak appears when the detector is surrounded by a dense material and is equal to 0.511 MeV.

**F: Backscattering Peak**

The backscattering peak is created when a Compton scattering even occurs outside the detector and the scattered photon reaches the detector and deposits its full energy. The scattering angle occurs over a small range of angles around 180°. This causes the peak to be a range of energies instead of a singular energy peak. The energy for backscattering peaks is usually between 200 and 300 KeV. For the case of 180° scatter, the energy of the incoming scattered photon would be

|  |  |  |
| --- | --- | --- |
|  |  | (2.6) |

**G: X-Ray Fluorescence Peaks**

X-ray fluorescence peaks are not of interest to prompt gamma neutron activation analysis and contributes to a collection of noise signatures that collect below 100 KeV. These peaks are generated by interactions that occur in the surrounding medium by exciting the target atom. During the de-excitation, characteristic X-ray photons are emitted and enter the detector and deposit their full energy. These peaks complicate the fitting functions, as they can be orders of magnitude higher than other signatures within the spectrum.

**H: X-Ray Escape Peaks**

When the excited atoms discussed above are generated within the detector, an X-ray photon can exit the detector. The resulting energy response is equal to the full energy peak minus the energy carried by the exited photon. If the detector material is larger than 1”, these events happen infrequently and do not affect the total response substantially.

**Energy Resolution**

Unfortunately, Figure 2-2 does not accurately represent the true response of even the highest resolution detectors. The energy resolution is dependent on the characteristics of the detector material and electronics. For scintillation detectors, the energy peaks are Gaussian distributed peaks and can be described by the normal distribution

|  |  |  |
| --- | --- | --- |
|  |  | (2.7) |

Where,

* + is the energy
  + is the peak centroid
  + is the standard deviation

The common way of determining the standard deviation is by measuring the full width at half maximum (FWHM), or the width of the peak at half of the amplitude. The standard deviation can be calculated by solving for

|  |  |  |
| --- | --- | --- |
|  |  | (2.8) |

It has been demonstrated (Wang and Gardner, 2012) that a subroutine can fit parameters to simplify the FWHM equation to

|  |  |  |
| --- | --- | --- |
|  |  | (2.8) |

Where,

* + is the energy
  + are empirical fitting parameters

**CHAPTER 3**

**Machine Learning Enhancements**

Many modern-day advancements have been made with the assistance of machine learning techniques. Traditional methods of solving linear least squares (LLS) problems can be enhanced by utilizing ready-made packages available on MATLAB and Python coding platforms. All codes used for this investigation have been modified from the sklearn packages in Python.

**3.1 Supervised Machine Learning**

**3.1.1 Linear Least Squares**

The linear model and analysis have been thoroughly used and examined over the last half century and remains important. The linear model, given a vector of inputs , an output Y can be predicted as

|  |  |  |
| --- | --- | --- |
|  |  | (3.1) |

Where,

* + is the intercept, also known as the bias in machine learning
  + is the predicted output

can be included in the vector coefficients and a constant variable 1 in X, allowing equation 3.1 to be rewritten as

|  |  |  |
| --- | --- | --- |
|  |  | (3.2) |

Where,

* + is the vector or matrix transpose
  + is the predicted output
  + is the linear coefficient
  + is the error

Fitting a linear model to a training data set is popularly done by selecting the coefficients that minimize the residual sum of squares

|  |  |  |
| --- | --- | --- |
|  |  | (3.3) |

The ordinary least squares estimator can be computed by using the following equation

|  |  |
| --- | --- |
|  | (3.4) |

**3.1.2 Least Absolute Selection and Shrinkage Operator (LASSO)**

The least absolute selection and shrinkage operator (LASSO) method became popular as a statistical and modeling method to reduce or eliminate unnecessary variables from a model (Tibshirani, 1996). The LASSO method utilizes tuning parameters to shrink and select variables placed in a linear model by reducing predictive error between in and out of sample tests, also known as test/train splitting in machine learning. LASSO is defined as

|  |  |  |
| --- | --- | --- |
|  | Or | (3.4) |

Where,

* + is the loss function
  + is a tuning parameter that serves as a penalty
    - Note: when , eq. 3.4 is identical to linear least squares (LLS)

**3.1.3 Elastic Net**

LASSO has been thoroughly investigated since its introduction. While offering many benefits over traditional linear least squares, limitations were found in situations where

1. In a case where the number of observations (n) are less than the number of parameters (p) LASSO can select at most n variables
2. In a case where there are a large number of predictors that are highly correlated, LASSO tends to select only 1, seemingly at random (grouping effect)
3. In a case where n>p and the predictors are highly correlated; LASSO is dominated by shrinkage methods like Ridge Regression

To remedy these deficiencies, Elastic Net was proposed (Zou and Hastie, 2005). The Elastic Net variable selection method combines the penalty terms used in LASSO and Ridge Regression as

|  |  |  |
| --- | --- | --- |
|  |  | (3.5) |

Where,

* + is the loss function
  + is a tuning parameter that serves as a penalty from LASSO
  + is a quadratic tuning parameter that serves as a penalty from Ridge Regression
    - Note: when , eq. 3.5 is identical to LASSO
    - Note: when , eq. 3.5 is identical to linear least squares (LLS)

**3.1.4 Coordinate Descent Solutions for LASSO and Elastic Net**

Before diving into the detailed derivation for coordinate descent solutions to LASSO and Elastic Net, some background information and notations are necessary (Gauraha, 2018). Consider the standard linear regression equation given as equation 3.2, assume the components of the noise vector are independent and identically distributed. Using subscripts j to denote the jth column of a dataset. Assuming that the design matrix **X** is fixed, the data is adjusted to be centered, and the predictors are standardized such that

|  |  |  |
| --- | --- | --- |
|  |  | (3.6) |

The -norms are defined as

|  |  |  |
| --- | --- | --- |
|  |  | (3.7) |

Then, the soft-thresholding operator can be defined as follows

|  |  |  |
| --- | --- | --- |
|  |  | (3.8) |

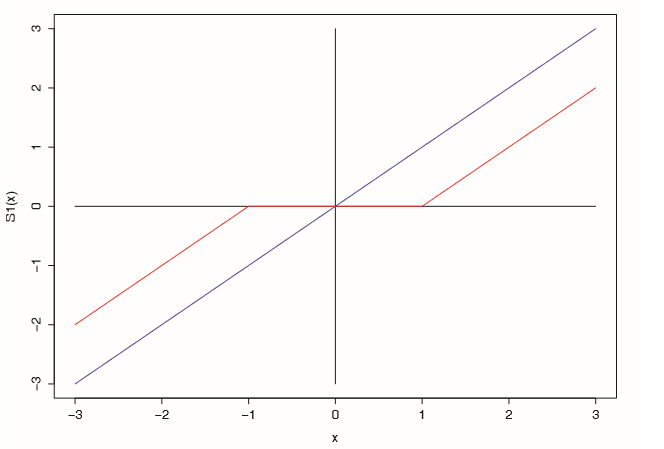


Figure 7: Soft-thresholding function

**3.1.4.1 LASSO Single Variable Case**

Before moving to a case involving multiple variables, a single variable case is considered. For this case, *p*=1 and , and the optimization problem can be written as

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Assuming is a solution to equation 3.9, then the sub-differential must contain zero, meaning

|  |  |
| --- | --- |
|  | (3.9) |

Which can be rewritten as

Note that since we are assuming the predictors are standardized, ,

|  |  |
| --- | --- |
|  | (3.9) |

An alternative interpretation is that is soft-thresholded by such that

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Therefore, the LASSO estimator for a single variable case can be computed by soft-thresholding the OLS estimator by or

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Where .

**3.1.4.2 Coordinate Descent for LASSO**

In order to treat LASSO algorithmically, the objective function must be split into a differentiable part and a non-differentiable part . The non-differentiable component is convex in each coordinate, allowing for a coordinate wise minimization to be utilized (Gauraha, 2018). Section 3.1.4.1 demonstrated that with a single predictor, the LASSO solution has a closed for solution with a soft-threshold version of the ordinary least squares estimate. Building on this, a coordinate descent algorithm for the LASSO can be implemented as follows.

Coordinate descent is an iterative method that solves one variable iteratively, while holding all other variables constant. For each coordinate sub-problem, each component of is fixed except for the jth component . By denoting as the jth column of ***X*** and denote all of the columns except for the jth column, then the problem can be rewritten as

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Next, we define as the partial residual or the difference between the actual response and the fitted model that excludes variable . The solution above becomes the univariate LASSO problem with vector as the response variable

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Now suppose is a solution to the optimization problem above. Then the stationary condition yields the following

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  |  |  | | (3.9) |

The OLS estimator for the jth variable can then be computed as The univariate LASSO solution can then be computed by soft-thresholding the OLS estimator as

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Table 3.1 demonstrates the full coordinate descent algorithm for LASSO. The dataset is loaded and initialized such that . Then for each j term, the single variable LASSO solution is computed. The entire process is repeated until .

Table 3.1: Full LASSO Coordinate Descent Algorithm

|  |
| --- |
| **LASSO Coordinate Descent Algorithm** |
| **Input:** dataset (**Y, X**) |
| **Output:** LASSO estimated vector of regression coefficients |
| Initialize |
| **repeat** |
| **for each**  **do** |
| Compute the partial residual , where |
|  |

Compute the OLS coefficient for single predictor

|  |
| --- |
|  |
| Update (LASSO solution: single variable case) |
|  |
| **end** |

**until** *convergence;*

|  |
| --- |
|  |
| **Return** |

**3.1.4.2 Coordinate Descent for Elastic Net**

The derivation and steps to calculate the Elastic Net penalty via coordinate descent are similar to those taken to calculate the LASSO penalty (Yang, 2013). The second order penalty adds an additional term to the LASSO solution such that

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Where is the Elastic Net penalty defined as

Note that when , the Elastic Net reduces to the LASSO equation. When each predictor shows a strong correlation, some should be used. For each fixed , coordinate descent is used to solve the Elastic Net. Once again, we define a current residual that excludes the jth term from each calculation. To update the estimate for , the univariate Elastic Net problem is solved by

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

The soft threshold is applied, leading to

|  |  |  |
| --- | --- | --- |
|  |  | (3.9) |

Table 3.2 demonstrates the full coordinate descent algorithm for Elastic Net. The dataset is loaded and initialized such that . Then for each j term, the single variable Elastic Net solution is computed. The entire process is repeated until .

Table 3.2: Full Elastic Net Coordinate Descent Algorithm

|  |
| --- |
| **Elastic Net Coordinate Descent Algorithm** |
| **Input:** dataset (**Y, X**) |
| **Output:** Elastic Net estimated vector of regression coefficients |
| Initialize |
| **repeat** |
| **for each**  **do** |
| Compute the partial residual , where |
|  |
| Update (Elastic Net solution: single variable case) |
|  |
| **end** |

**until** *convergence;*

|  |
| --- |
|  |
| **Return** |

**3.1.5 Cross Validation**

Cross validation is a method to perform out of sample testing to assess the results of statistical analysis on an independent data set. After the experimental and simulated data are processed and separated into **X** and **Y** vectors, the full data set is split into testing and training subsets via the holdout method shown in figure 3.?. As a result of splitting the data, additional bias is introduced to each solution, as different parts of the data are either included or removed from the model selection process. Cross validation is used to perform the test-train split process multiple times, averaging each solution to provide a less biased solution as seen in figure 3.?. Cross validation is used to find the optimal normalization parameter for both LASSO and Elastic Net.

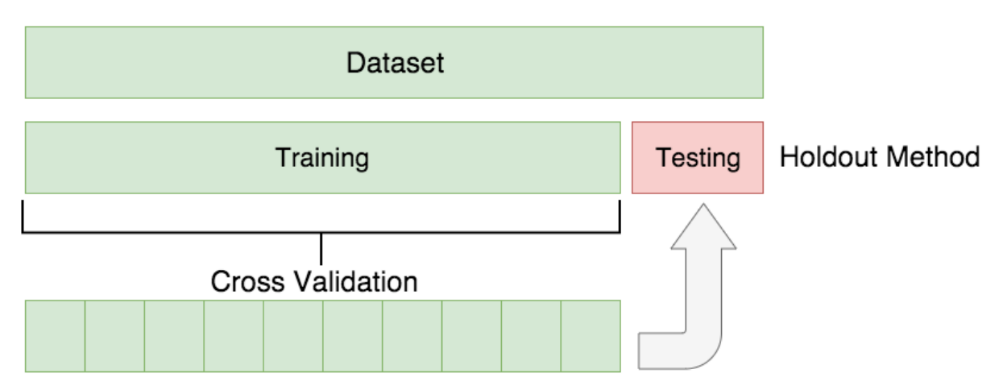


Figure 8: Holdout method for test-train split

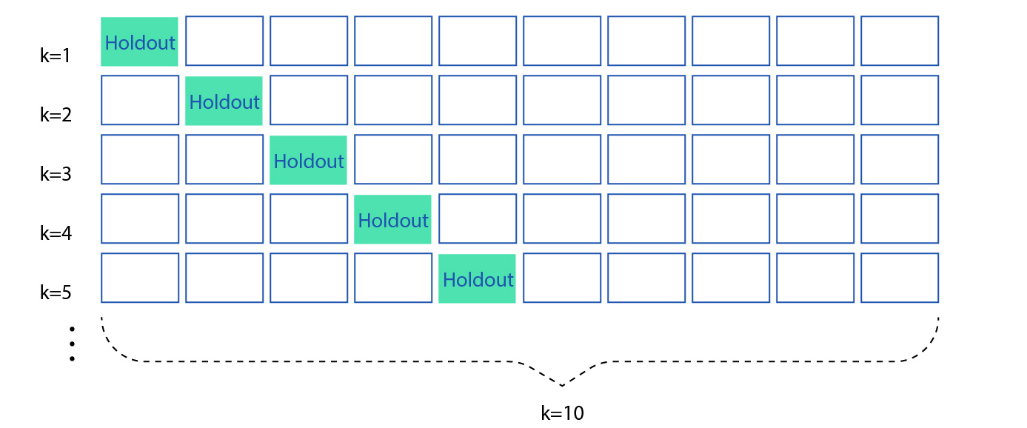


Figure 9: Cross validation holdout

**CHAPTER 4**

**Kansas State Experimental Data Processing and Results**

**4.1 Data Collection**

**4.1.1 Kansas State Experimental Data**

In section 1.2, the Kansas State University benchmarking tool and experimental facility were described. Five separate experimental runs were conducted for tap water, pure sand, sand with water, pure limestone, and limestone with water. For each of these experimental runs, the data collection sequence involved:

1. Source calibration runs to adjust gain on detection systems
2. Pre-run background tests (5 minutes)
3. Live D-T run (1 hour)
4. Post-run background tests (5 minutes)

The data collected during each of these runs was processed by Kansas State University and converted into counts per channel text files for each detector. Additional binary files were provided that have time dependent information for the pulsing sequence and detector responses with respect to time.

The counts per channel data files must be processed and converted to energy to match the simulated responses. A second order polynomial is used for fitting purposes by

|  |  |  |
| --- | --- | --- |
|  |  | (4.1) |

**4.1.2 Simulated Data**

In order to simulate an accurate detector response, extensive (~109 particles) MCNP 6.1 calculations were conducted with accurate details of the detector, tool, test chamber, and surrounding medium geometry and materials. Libraries were created by using F8 tallies for each detector and medium (water, sand, limestone, etc.). To accurately model the detector response, Gaussian broadening functions were used in order to take the pulse response and broaden the data to match the FWHM found using the calibration sources.

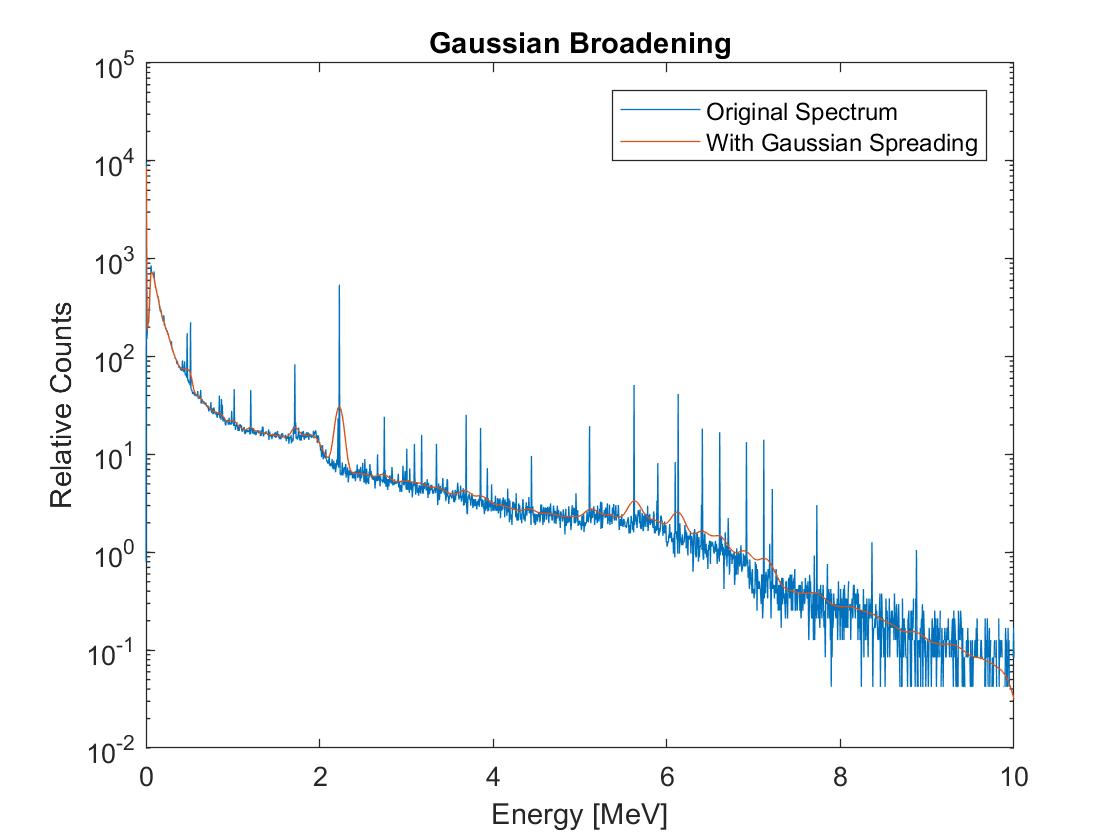


Figure 4-1: F8 tally simulated response and Gaussian broadened detector response

**4.1.3 Simulated Example**

Prior to running a full example with data from Kansas State University, a simulated example was conducted to demonstrate the capabilities of each code. Using simulated data, a salt water test case was created by combining a combination of water, sodium, and chlorine libraries. Random Poisson distributed noise was then applied to the test case to simulate the random effects found in true examples. Five total libraries were used as training data sets (water, sodium, chlorine, iron, and copper) to train the model versus the test data set containing salt water. Using 10-fold cross validation, the model was correctly trained to fit only the 3 libraries that contribute to the full spectrum, while providing a zero contribution for the iron and copper libraries. Figures 4-2 thru 4-4 demonstrate how the tuning parameters select the best model and the final fits for LASSO and Elastic Net.

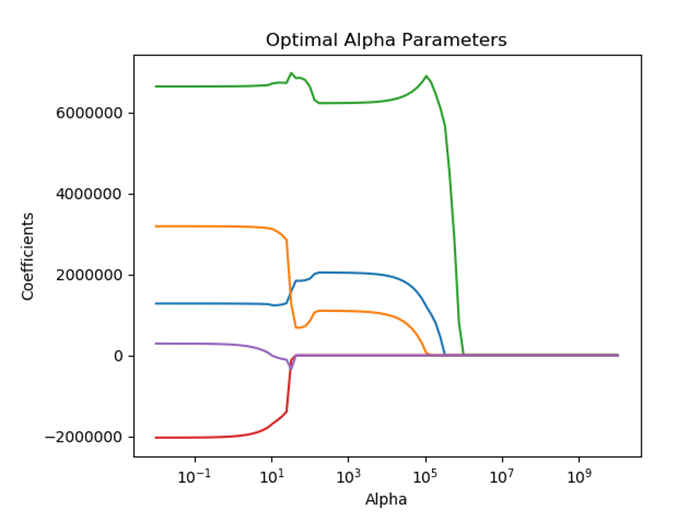


Figure 4-2: Model selection by changing tuning parameters

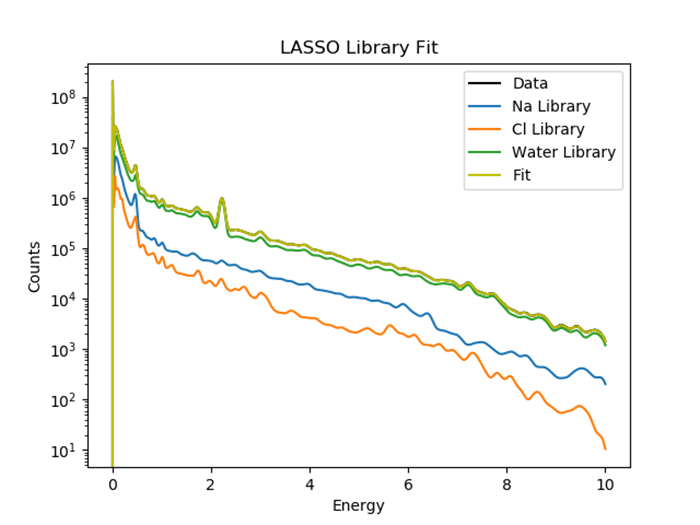
****

Figure 4-3: LASSO salt water simulation fit

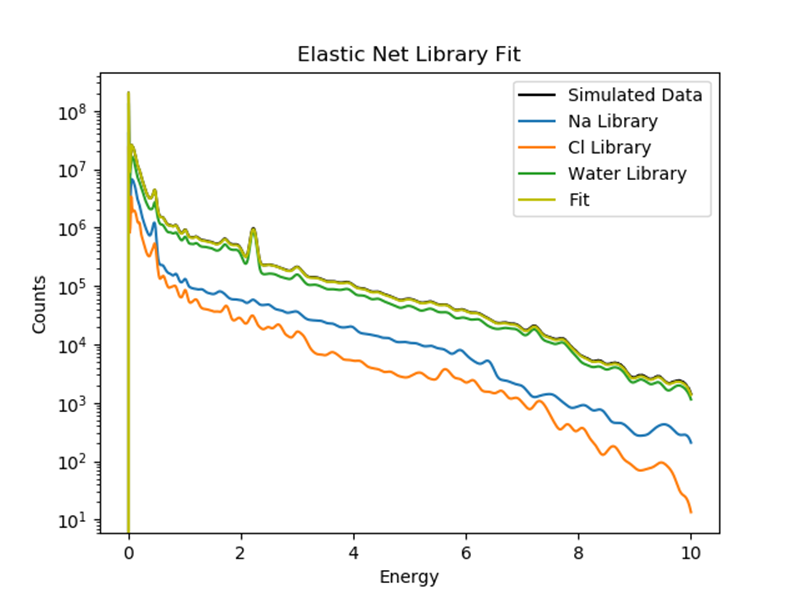
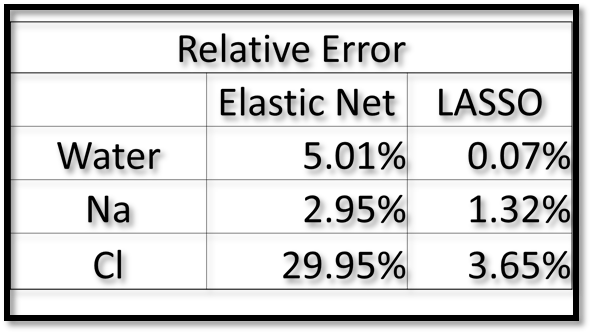


Figure 4-4: Elastic Net salt water simulation fit

When the normalization parameter is increased, the penalty for each term becomes greater, allowing the shrinkage and removal of parameters. The MSE at each alpha parameter is compared, settling on the best overall fit for the final solution. Without putting the output of the LASSO and Elastic Net fits into an OLS program for a final fit, table 4-1 displays the relative error of each method averaged over 10 runs.

Table 4-1: Relative error of the two methods



**4.2 Methods and Improvements**

**4.2.1 Full Procedure**

Before any data analysis can take place, a variety of preprocessing and post processing of data is necessary (Fig. 4-5). Experimental and simulated data are generated at Kansas State and North Carolina State, respectively. The experimental data is processed and separated into counts per channel text files for each detector type, as well as binary files with time dependent data. The detector data is converted into energy bins using calibration sources for empirical FWHM calculations.

The simulated data is generated using MCNP 6.1 using the F8 tally. The nonlinear response in the Gaussian broadening detector response function is addressed, resulting in a final library spectrum.

Each library is added to a single file and used to train the fitting model. A 10-fold cross validation process helps reduce the overall bias in model selection. For the test/train split, 90% of the data is used each time for training, while 10% of the data is held out to test the model. LASSO and Elastic Net are applied to select the best model that reduces a loss function (MSE). Once the final solution is reached, the output from LASSO and Elastic Net are used as initial guesses in an ordinary least squares fitting using CEARLLS.

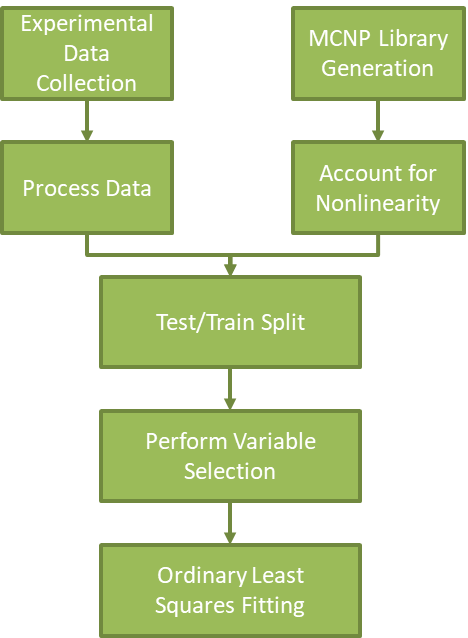


Figure 4-5: Full procedure

**4.2.2 Early Lessons Learned**

A pure water trial was conducted and processed by the procedure laid out in section 4.2.1. Upon completing the final fitting, several deviations were observed as seen in Figure 4-6. In the high energy range, above 7 MeV, differences between the simulated and experimental response are expected, as the error in each channel is high from the low count rate. The hydrogen peak and oxygen peaks fit well to the simulations, however, the region below 2 MeV does not have a proper fitting. The proposed solution was to look at the activation of the NaI crystal (Gardner, 2000).

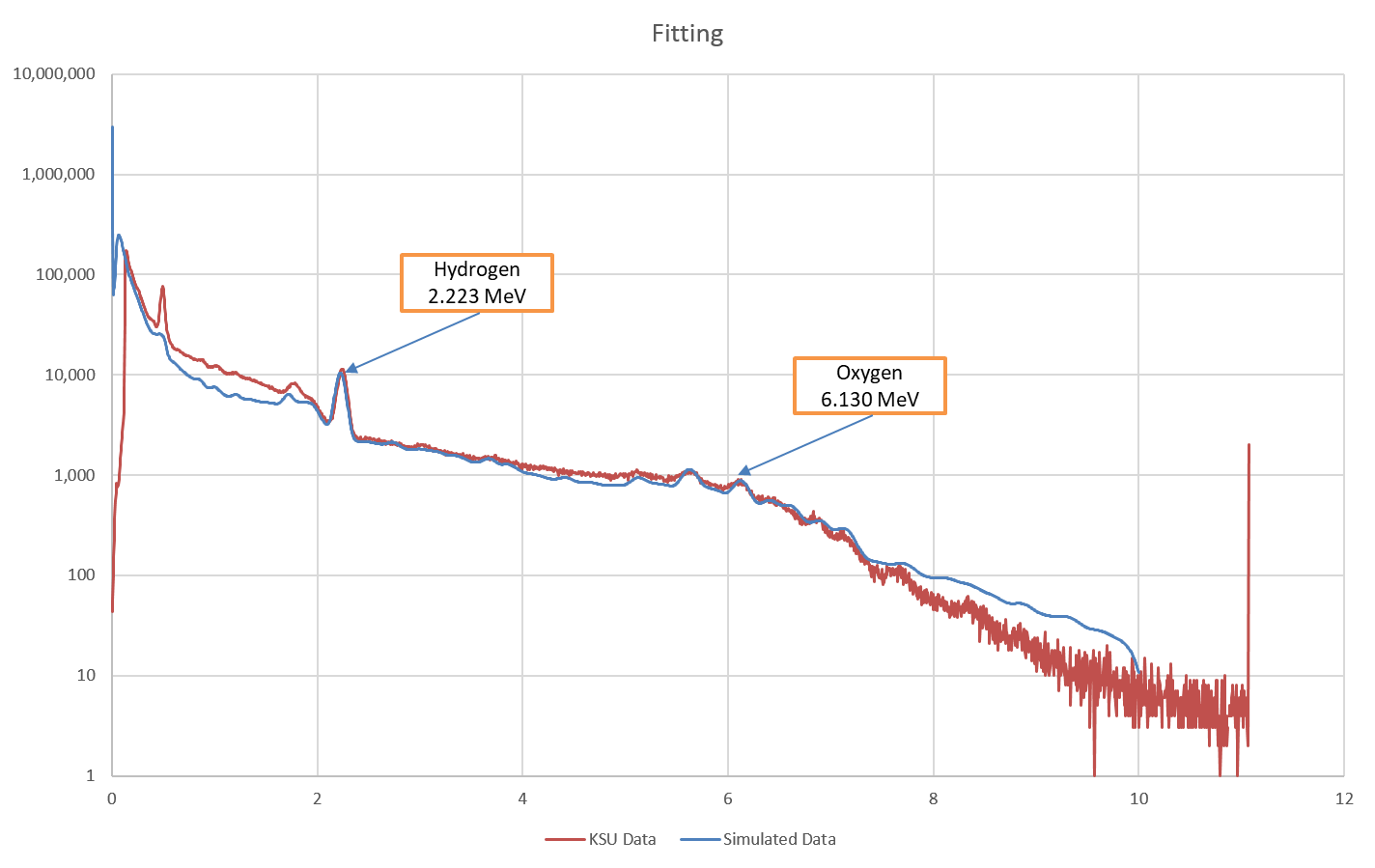


Figure 4-6: Pure water first fit

To test this theory, a library response for activated Sodium and Iodine is added to the fitting libraries (Fig. 4-7). Although this did improve the overall fit, it did not completely resolve the discrepancies.

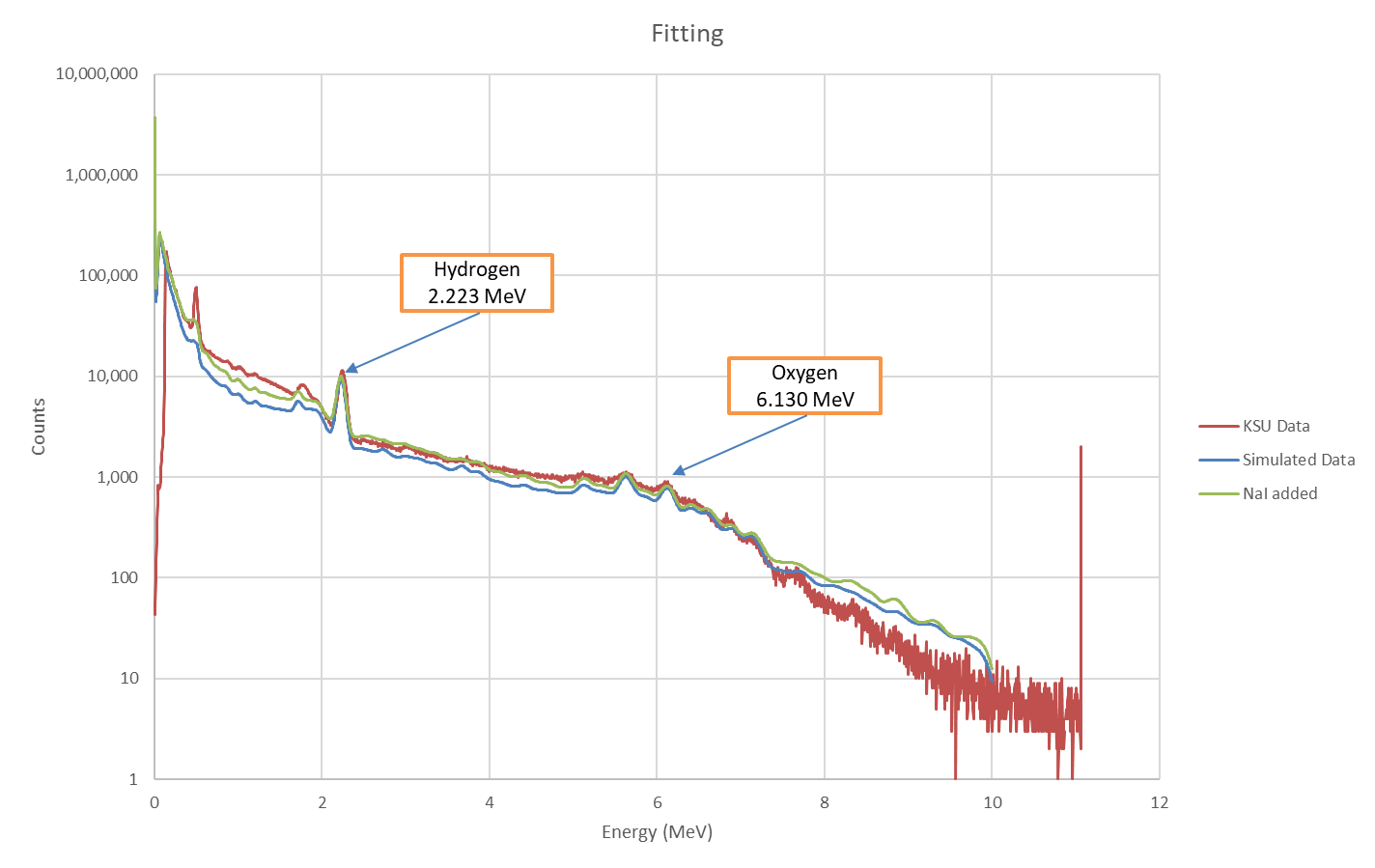


Figure 4-7: Pure water fit with NaI activation libraries added

To completely understand the nature of the activated (delayed) response, the time dependent DT data is necessary. The primary advantage of using a PNG is the ability to read signatures that are prompt and delayed by the response times. The PNG is triggered using a firing sequence from the generator and a pulse of neutrons is emitted. After the pulse, there is a window of time before the next initialization where the neutrons have died off, and the only remaining signatures are from delayed activation (Figs. 4-8, 4-9).

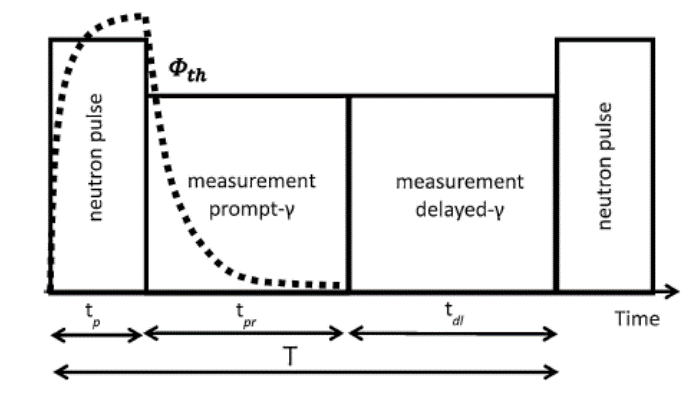


Figure 4-8: Firing sequence and measured response

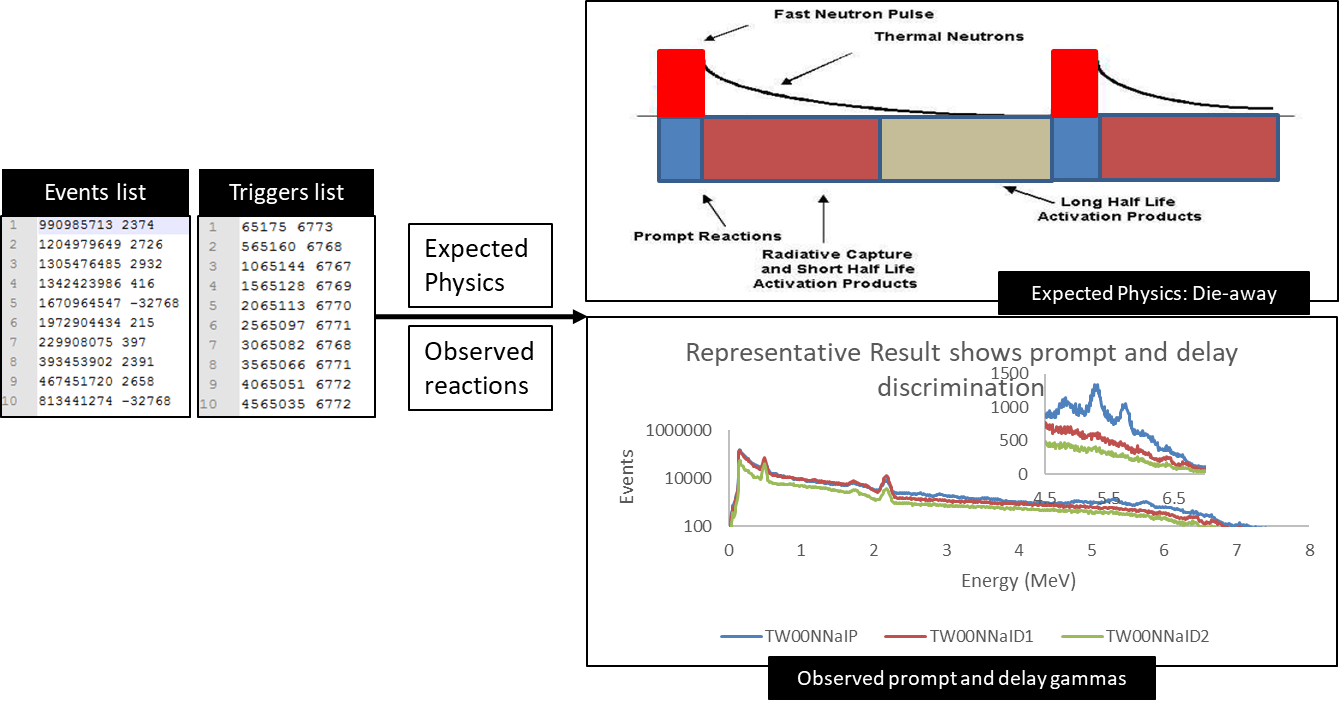


Figure 4-9: Time dependent format distributed by Kansas State University

**4.3 Kansas State Experimental Results**

**4.3.1 Water Trial**

The water trial (Fig. 4-10) shows a marked improvement in the fitting for the low energy region. Both near and far detectors show a proper fit for the Oxygen and Hydrogen peaks and all secondary features. Final fitting and analysis will be performed upon improvements made in the discussion section.

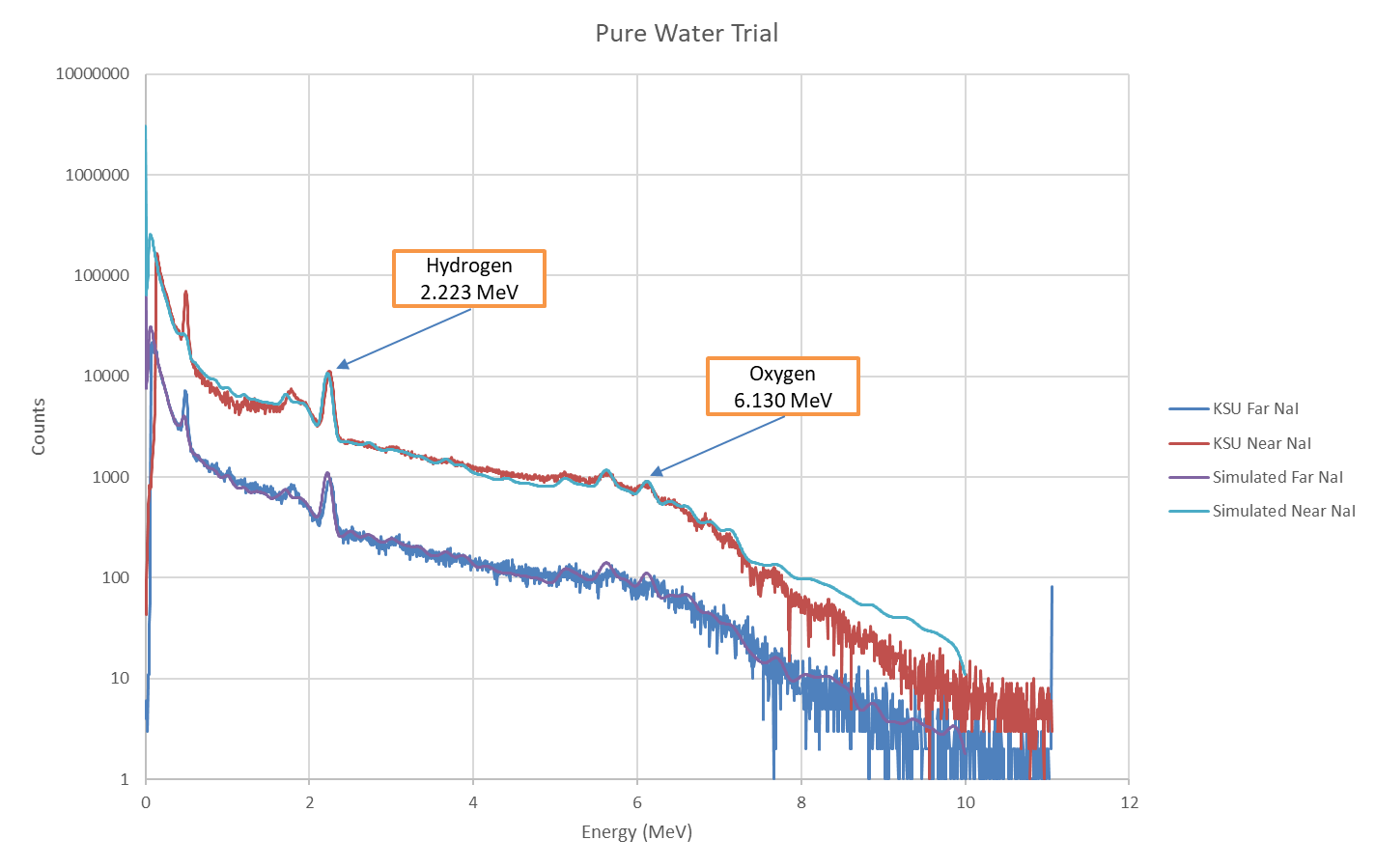


Figure 4-13: Near and far detector fit after removing background

**4.3.2 Sand Trial**

The sand trial (Fig. 4-11) shows proper fitting for the near detector. The far detector shows a slight altering of the energy fitting, likely due to a shift in gain parameters. Final fitting and analysis will be performed upon improvements made in the discussion section.

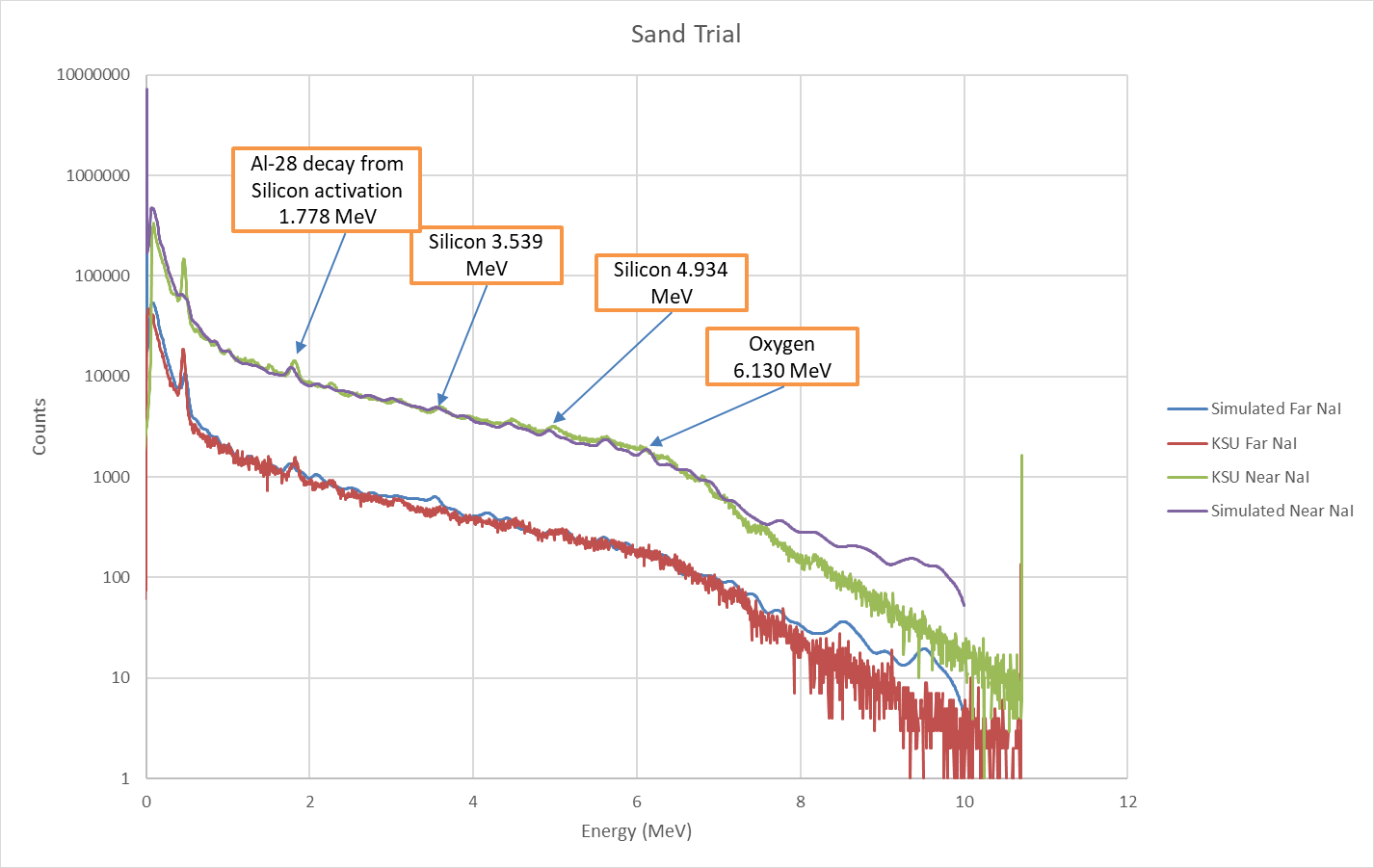


Figure 4-14: Near and far detector fit after removing background

**4.4 Discussion/Next Steps**

The LASSO and Elastic Net variable selection techniques perform perfectly for both the sand and water trials. Each data set was run 20 times, and each time, the correct libraries were selected, while excluding libraries that did not contribute to the spectrum. Improvements continue to be made by both the NC State group and the Kansas State group. Aaron Feinberg is working on a Bayesian approach to fit the non-linear components instead of the basic calculations currently being used. Long Vo is working on separating the time dependent data to improve the background contribution.

Additional data sets for limestone, as well as some mystery data sets will be sent shortly. These data sets will challenge these methods to ensure that we are accurately predicting the correct compositions. The mystery data sets will have multiple materials in each set and should offer a greater challenge on the limitations of each method.

The final investigation to conduct is on the limitations of LASSO and Elastic Net in selecting variables in radiation detection applications. To do this, handheld RIID simulations will be conducted to simulate many different scenarios. The number of radioisotopes in each sample, the number of total channels, and the number of total counts will be varied in order to test the detection limits of each method. The final challenge will be to introduce shielding to the problem and demonstrate the effectiveness in a more difficult test.

**Bibliography**

A.M. Yacout, R.P. Gardner, K. Verghese, Nucl. Instr. And Meth. A 243 (1986) 121.

A.Sood, R.P.Gardner / Nucl. Instr.and Meth.in Phys. Res.B 213 (2004) 100–104.

Bob Burkhart, “Neutron Generator and Well Logging”, Sandia National Laboratories, SAND2006-3291C, [www.osti.gov/servlets/purl/1264595](http://www.osti.gov/servlets/purl/1264595)

Committee on Radiation Source and Replacement, 2008, Radiation Source Use and Replacement, National Research Council, ISBN 0-309-110115-7, National Academies Press, Washington, DC (NAS-2008)

E. Bai, K. Chan, W. Eichinger, P. Kump, “Detection of radionuclides from weak and poorly resolved spectra using Lasso and subsampling techniques.” Radiation Measurements, 46 (2011), pp. 1138-1146.

Evans, R.D., *The Atomic Nucleus.* McGraw-Hill, 1955.

Gardner, et al., 2013, ANS Panel on “Recent Developments in Radiation Source Use and Replacement after the NAS Report of 2008”, Transactions of the American Nuclear Society, Vol. 109, 47.

R.P. Gardner and A. Sood. “A Monte Carlo simulation approach for generating NaI detector response functions (DRFs) that accounts for non-linearity and variable flat continua”. *Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms* 2013 \*Jan. 2004), pp. 87-99.

R.P. Gardner and C.W. Mayo. “Nai detector nonlinearity for PGNAA applications”. *Applied Radiation and Isotopes* 51.2 (Aug. 1999), pp. 189-195.

Han, Xiaogang, Ph.D thesis. North Carolina State University. 2005.

Hastie, Trevor. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Edition. Springer Series in Statistics. 2009.

Hou, Guojing, Ph.D thesis. North Carolina State University. 2017.

Knoll, G. F., *Radiation Detection and Measurement*. 4th Edition. John Wiley and Sons, Inc, 2011.

Liberman, A.D et. al., “A method to determine the absolute neutron output of small D-T generator”, *Nuclear Instruments and Methods in Physics Research Section B,* Volume 79, Issue 1-4, p. 574-578, June 1993

M. Geretschlager, Nucl. Instr. and Meth. B 28 (1987) 289.

MCNP Training Manual

Preston R.M. et al, “Neutron generator burst timing measured using a pulse shape discrimination plastic scintillator with silicon photomultiplier readout”, *Journal of Instrumentation*, December 8, 2013

R.P. Gardner, A.M. Yacout, J. Zhang, K. Verghese, Nucl. Instr. and Meth. A 242 (1986) 399.

R.L. Heath, Scintillation Spectrometry, Gamma-Ray Spectrum Catalogue, 2nd Ed., USAEC Report IDO-16880, 1964.

T. Papp, J.L. Campbell, D. Varga, G. Kalinka, Nucl. Instr. and Meth. A 412 (1998) 109.

Serra, Oberto. *Well Logging Handbook.* Ed. Technip, Paris*.* 2008.

Tibshirani, R. “Regression shrinkage and selection via the lasso.” *Journal of the Royal Statistical Society, Series B*  58: 267-288. 1996.

Y. Jin, R.P. Gardner, K. Verghese, Nucl. Instr. and Meth. A 242 (1986) 416.

Zou, H. and Hastie, T. “Regularization and variable selection via the elastic net”. *Journal of the Royal Statistical Society Series B.*  67: 301-320. 2005.

Burr, T and Hamada, M. “Radio-Isotope Identification Algorithms for NaI Gamma Spectra.” 2009

Blackadar, J. et at. “Evaluation of Commercial Detectors. 2003

Pribida, L et al. “Evaluation of Handheld Radionuclide Identifiers. “ 2004