3-2022

Characterization of Environmental Conditioning of Lithium Hydride Using Spectroscopy and Machine Learning

Ryan E. Pinson

Characterization of Environmental Conditioning of Lithium Hydride Using Spectroscopy and Machine Learning

THESIS

Ryan E Pinson, B.S.P., 1st Lt, USAF

AFIT-ENP-MS-22-M-105

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENP-MS-22-M-105

Characterization of Environmental Conditioning of Lithium Hydride Using Spectroscopy and Machine Learning

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Applied Physics

Ryan E Pinson, B.S.P.,

1st Lt, USAF

March 22, 2022

AFIT-ENP-MS-22-M-105

Characterization of Environmental Conditioning of Lithium Hydride Using Spectroscopy and Machine Learning

THESIS

Ryan E Pinson, B.S.P.,
1st Lt, USAF

Committee Membership:

Anil K Patnaik, Ph.D
Chair

Christina L Dugan, LTC, Ph.D
Member

Phillip R Jenkins, Maj, Ph.D
Member

Andrew V Giminaro, Ph.D
Member

AFIT-ENP-MS-22-M-105

Abstract

Lithium compounds such as lithium hydride (LiH) and anhydrous lithium hydroxide ($LiOH$) have various applications in industry but are highly reactive when exposed to moisture and CO_2. These reactions create new molecular forms, including compounds such as lithium oxide (Li_2O), lithium hydroxide monohydrate ($LiOH \cdot H_2O$), and lithium carbonate (Li_2CO_3). These new compounds degrade the effectiveness in applications using these compounds. The negative effects induced by new lithium compounds creates a need for the ability to characterize the in-growth of such compounds. To study these in-growths, this work will present environmental conditions such as heat, moisture, and the atmospheric conditions, as examples of storage conditions. A pulsed laser and an echelle spectrograph are used in a novel single setup to conduct both Raman spectroscopy and Laser-Induced Breakdown Spectroscopy (LIBS) in tandem. By employing spectroscopic techniques such as LIBS and Raman spectroscopy, in conjunction with multivariate modeling techniques (PCA,PCR,PLSR,Random Forest), these various conditions will be explored. These measurements and analysis techniques will enable collection of critical information required for validation of modeling on environmental characterizations of the lithium based compounds and their reactions that have significant implications on industrial technologies, such as batteries, and nuclear security applications.

Table of Contents

Page

List of Figures

Figure Page

Figure Page

Figure Page

List of Tables

Characterization of Environmental Conditioning of Lithium Hydride Using Spectroscopy and Machine Learning

I. Introduction

Lithium compounds have been used in industries ranging from pharmaceuticals, to batteries, to nuclear weapons, and in a variety of other manners [10, 11, 12, 13]. Specifically, LiH and the reactions and compounds that it can undergo are of interest. LiH reacts with moisture to form $LiOH$, and then $LiOH$ can form Li_2CO_3 when exposed to a source of carbon dioxide [14, 15]. These reactions can lead to a in-growth of alternate forms of lithium compounds that degrade the usefulness and performance for the intended application. The quantification of the degradation present in such lithium compounds is thus desired by industries and other stakeholders.

This work employs two spectroscopic measurement techniques and uses various analytical techniques to analyze the spectral measurements. The two measurement techniques used in this study are Laser-Induced Breakdown Spectroscopy (LIBS) and Raman Spectroscopy. The analytical approaches are multivariate analysis techniques such as Principal Component Analysis (PCA), Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR), and a decision tree model known as Random Forest. Lithium carbonate as it relates to lithium hydride and hydroxide has not been studied to the same extent that the hydrolysis of lithium hydride to form lithium hydroxide has been [7, 8, 9, 16, 17, 18, 19, 20, 21, 22, 23, 24]. The reactions that form lithium hyroxide monohydrate and lithium carbonate will be of particular interest in this experiment. While this experiment does not focus heavily of the chemistry of these reactions, they are important to understand for interpreting

the spectral responses in this experiment.

1.1 Motivation

Lithium hydride is used in a variety of applications from everyday objects, such as batteries, to highly controlled devices such as nuclear weapons [16]. For examples, LiH rapidly provides gaseous hydrogen upon contact with water, which has been studied for use in life preserving buoyancy devices [11]. Also, hypersonic flight has recently been a important research topic, and LiH has shown great promise as a coolant in such devices [25]. Lithium hydride has also been used as a moderator in nuclear reactors, along with depleted uranium [10]. The primary challenge with LiH is the ease at which it reacts and forms $LiOH$, Li_2O and Li_2CO_3 [14, 15]. Since the 1950s, Li_2O and $LiOH$ have been studied extensively as related to LiH, but Li_2CO_3 has not had the same level of study applied to it [16, 20, 22, 26, 27, 28, 29]. To enhance optimization of storage and use conditions, additional studies of lithium carbonate formation from lithium hydride, lithium hydroxide, and lithium oxide could be developed. However, these studies would need forensics analysis, which can be difficult to obtain, especially outside of a laboratory setting. This work hopes to present data collection and analysis techniques that could be used to improve the capability of lithium compound characterization. In addition, this type of characterization could be extended to various applications in industry and DOD environments.

1.2 Background

This section will discuss the the chemical reactions involved in this work and the background of the methods employed in this thesis.

1.2.1 Lithium Compounds

As stated before, lithium hydroxide and lithium oxide formation from lithium hydride has been studied repeatedly since the 1950s [16, 20, 22, 26, 27, 28, 29]. LiH is the lightest solid hydrogen storage material known. That makes it a prime candidate for many applications of hydrogen storage where weight is of concern (e.g. space exploration). However, overcoming the reactive nature with moisture in the environment limits it application usefulness. This is one reason the reactions between LiH and moisture have garnered interest. Li_2CO_3 formation reactions are noted in multiple studies in the context of surface barriers and as a natural contaminant [28, 29]. However, its formation is not well characterized.

The reaction known to produce lithium hydroxide is described by the following equations. It has been observed that at the boundary layer between the $LiOH$ and LiH the reaction shown in Eqn 2 occurs. The formation of lithium carbonate follows the reactions in Eqn 3 and 4.

$$LiH + H_2O \rightarrow LiOH + H_2 \tag{1}$$

$$LiOH + LiH \rightarrow Li_2O + H_2 \tag{2}$$

$$Li_2O + CO_2 \rightarrow Li_2CO_3 \tag{3}$$

$$2LiOH + CO_2 \rightarrow Li_2CO_3 + H_2O \tag{4}$$

While Eqn 4 is technically correct it actually skips and intermediate step which is valuable in distinguishing certain stages of the reaction. The true reactant with CO_2

is $LiOH \cdot H_20$ or H_3LiO_2, known as lithium hydroxide monohydrate. This being the case, the overall reaction can therefore be written as two steps [30]:

$$2LiOH + H_2O \rightleftharpoons 2LiOH \cdot H_2O \; or \; 2H_3LiO_2 \tag{5}$$

$$2LiOH \cdot H_2O + CO_2 \rightarrow Li_2CO_3 + 3H_2O \tag{6}$$

Thus, for lithium hydride to have an in-growth of lithium carbonate in a sample, the sample must first be exposed to moisture [29]. These various reactions provide necessary compounds to be aware of when attempting to determine characteristics within a lithium compound.

1.2.2 Raman Spectroscopy

Raman spectrocopy is a technique that is based on scattering of light and the Stokes shifts that occur based light incident on compounds. The Raman technique relies on off-resonant, nonlinear scattering to determine molecular composition of a sample. Rayleigh scattering, which is a type of elastic scattering, is the dominant scattering mechanism for any incident light while Raman scattering is much less intense. To deal with excessive light saturating the camera, Raman spectroscopy requires long integration times and special light rejection techniques to allow for a view of the Raman spectral lines. The spectral lines of Raman scattering change based on the excitation source. Otherwise the composition of the samples are easily characterized based on the scattered light's wavelength. This light is directed, using optical mirrors, lenses, and fibers, into a spectrometer which separates the spectral lines by wavelength in physical space. The intensity of each wavelength is measured by a camera and thus a spectra is recorded. This is discussed in further detail in

Section 2.1.

1.2.3 Laser Induced Breakdown Spectroscopy (LIBS)

Laser induced breakdown spectroscopy is a technique which uses a pulsed laser focused on a sample of interest. The laser ablates a portion of the sample, generating a laser induced plasma (LIP). As this super-heated plasma cools it emits light based on elements present in the plasma, determined by the compound incident upon. This light is directed, using optical mirrors, lenses, and fibers, into a spectrometer which separates the spectral lines by wavelength passing through the grating in the spectrometer. The intensity of each wavelength is measured by a camera and thus a spectra is recorded. Due, to the atomic nature of the spectral lines, which could be from multiple sources, a ratio of these lines to one another can be used to determine relative concentrations of each element present in the sample.

1.2.4 Chemometrics

Chemometrics is "the chemical discipline that uses mathematical and statistical methods to provide maximum chemical information by analyzing chemical data" [31]. The statistical methods employed in extracting information from the LIBS and Raman spectra in this study are PCA, PCR, PLSR, and decision tree models. These techniques are used to analyze the data and build models for predicting various information about samples. The results of these chemometric techniques will be used to indicate various results or determined characteristics in samples.

Chemometrics in combination with LIBS has been shown to be a valuable process to characterize samples [32, 33]. Using molecular information, similar to that gained from Raman spectroscopy, predictive models have been created as well. These models have been used to provided analysis for samples, using the fact that the models can

account for the molecular data [32, 33, 34, 35, 6]. By using the fact that various combinations of molecular compositions will have corresponding spectra or certain ratios of atomic emission peaks indicate a mixture, models can be used to predict sample composition.

1.3 Research Objectives

This research hopes to quantify and predict characteristics of samples following various environmental exposures. While studying the characterization of lithium samples, the capability of Raman spectroscopy, LIBS or a data fusion of the two will also be evaluated. In addition, differing models and their ability to evaluate and model the data will be investigated. A special interest is placed on humidity and temperature in this study and the varying responses LiH has when exposed to different conditions.

Previously, work was done providing evidence that Raman spectroscopy and LIBS are able to distinguish between various lithium compounds [36]. Thus, it is hypothesized that both LIBS and Raman will effectively characterize lithium compounds in pressed pellets for the environmental conditions in this study. This is also based on the aforementioned successful use of chemometrics which demonstrated that LIBS can be used for molecular composition analysis [33, 6]. By its very nature, Raman spectroscopy provides a spectra that is molecule specific [37]. Raman spectroscopy will likely provide a much better model (higher confidence, lower uncertainty) than LIBS for lithium compound characterization. Using multivariate models and machine learning for a given known environmental condition of a LiH exposure in training data, samples can test these models ability to quantify and predict characteristics of environments.

1.4 Experimental Process Overview

The parent lithium compound LiH was exposed to various environments containing H_2O and CO_2, along with various temperatures. These environmental treatments were performed on LiH, then the products were pressed into pellets for analysis. Next, Raman spectroscopy and LIBS were used to analyze the pellets to form spectral data sets. Raman spectroscopy was performed before LIBS, because it is imperative the spectra be taken from a pristine sample. The ablation process perturbs the surface of the samples which would then influence further measurements. Since Raman spectroscopy relies purely on scattering light, the samples are unaltered by performing this technique. After Raman, LIBS was performed on the samples. Special care was taken to ensure each position under analysis was unperturbed by any other ablation site. Finally, the LIBS and Raman data were analyzed with the multivariate techniques: PCA, PCR, PLSR, and Random Forest. Using these techniques, predictive models were built and tested in their characterization capability to determine information about various samples.

II. Theory and Literature Review

This section will discuss the relative information concerning Raman spectroscopy, LIBS, the study of lithium compounds with relation to both, chemometrics, and the data manipulation and analysis techniques used in this experiment.

2.1 Raman Spectroscopy

Although Raman spectroscopy has been used in recent works to study lithium hydride hydrolysis products to include $LiOH$, LiD, Li_2O, and $LiOH \cdot H_2O$, the present work explores effects of environmental conditions on lithium compounds [20, 8, 38]. Raman spectroscopy is a method that can identify the molecules present in a sample. This is accomplished by observing the scattered light from molecular vibrations and rotations following a monochromatic light source excitation. Raman spectroscopy is a non destructive method using scattered light from the sample.

Raman scattering was discovered by Sir Chandrasekhara Venkata Raman. C. V. Raman discovered when light is incident on a material, the majority of the light interacts by elastic scattering while a small portion of that light interacts through inelastic scattering. The scattered light does not retain the same wavelength or intensity as the incident light [39, 40]. The scattering is considered to be Rayleigh scattering is when the scattered light is of the same wavelength and energy as the incident light. Raman scattering is the inelastic scattering in which a nonlinear optical process occurs when the incident light interacts with a molecular state. After the interaction with the light, the molecular state transitions it to a new molecular state via a virtual state-involving absorption of the incident light and emission of light of different colors. The process is nearly instantaneous and occurs on the same timescale as the dominant Rayleigh scatter.

For Raman scattering, the rotational-vibrational (ro-vibronic) state of the molecule is excited. This occurs through a transition to a virtual energy state followed by a transition to another molecular ro-vibronic state, thus emitting a blue-shifted or red-shifted light. These transitions are called anti-Stokes and Stokes Raman scattering, respectively. Anti-Stokes emission lines will have a higher energy than the incident light, and thus a shorter wavelength. The opposite will be true for Stokes lines [1]. A visual representation for Rayleigh, Stokes Raman, and anti-Stokes scattering is seen in Figure 1. Although many varieties of molecular vibrations that could take place in typical Raman scattering, a summarized list of the fundamental vibrations is listed below:

- stretching: a change in the length of the bond
- bending: a change in the angle between two bonds
- rocking: a change in angle between a group of atoms and the rest of the molecule
- wagging: a change in the plane between the plane of a group of molecules and the plane of the rest of the molecule
- twisting: a change in the angle between the planes of two groups of atoms
- out-of-plane: the atom moves in and out of the plane of the other atoms

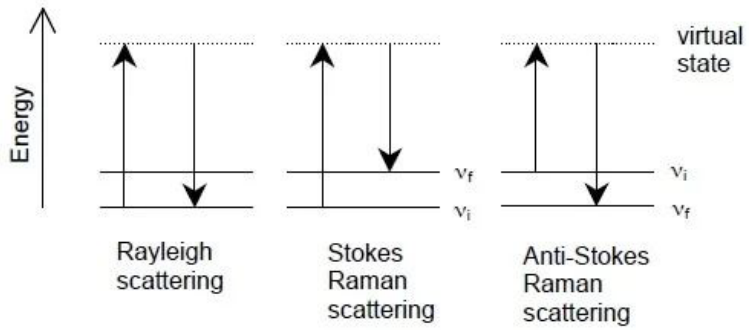

Figure 1: Rayleigh, Stokes, and Anti-Stokes Scattering Diagram [1]

Raman scattering typically occurs approximately once in every ten million scatters [39]. This makes detection or these scatters increasingly dependant of incident light intensity. Any device attempting to observe and measure these rarer Raman emissions would be saturated by the Rayleigh scatter long before the Raman lines come into the limits of detection [39]. For this reason, special notch filters that reject light of select wavelengths are needed in the optical setup for Raman spectroscopy. This filtering allows Raman lines to be viewed with common detection cameras such as an Intensified Charge-Coupled Device (ICCD). An example of a typical Raman spectroscopy setup is shown in Figure 2

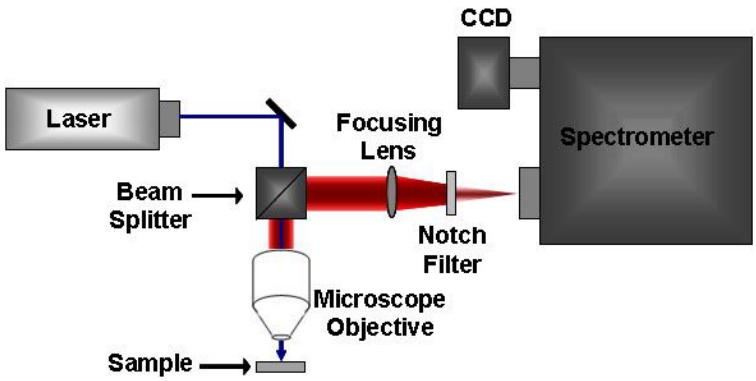

Figure 2: Example of Raman Setup [2]

Traditional commercial Raman spectroscopy systems use a continuous wave laser at low intensity. In order to achieve good signal, the sample and microscope objective must be kept in isolation from any stray light, as it would interfere with the spectra. Recently a novel setup was proposed by George and Shameem using a pulsed laser coupled with an echelle spectrograph [41, 42]. The camera on the echelle spectrograph is gated, and the relative intensity of the Raman scattered light, compared to the room, is very high in the time frame of the pulse duration. The ambient light from the room does not interfere, and thus eliminating the need for special dark rooms [42]. This newer approach also allows the coupling of both LIBS and Raman in one

experimental setup.

2.2 Laser Induced Breakdown Spectroscopy (LIBS)

LIBS is based on atomic spectral lines and a plasma's formation and cool. Observation of atomic emission can be traced back as early as the 1500s where the it was noted the color of flames changed with the introduction of different materials, specifically in the smelting of metal ores where the change in flame color indicated when the metal was ready to be cast [43]. Over the years, the use and precision of atomic emission spectra advanced until plasma's were used for atomic emission spectroscopy in 1964.

For LIBS to work, a large energy fluence is needed on a sample. This is accomplished using a pulsed laser and optical lenses to focus a laser in both time and space. When sufficient energy fluence is incident on a sample, the material becomes super-heated and creates a plasma with ions, electrons and dissociated atoms. As the super-heated plasma cools, a valence electron in a higher energy state returns to a lower energy state in the plasma, and a photon is emitted with the corresponding energy [43]. These characteristic photons are captured, and their specific wavelengths are the associated atomic spectral lines. Raman scattering which generates signal only during laser excitation, however LIBS have a time evolution associated with the plasma cooling that is paramount to understanding the timing needed for capturing meaningful data. Th mechanisms involved in LIBS are not necessarily simple, but the optical setup has relatively few parts as displayed in Figure 3.

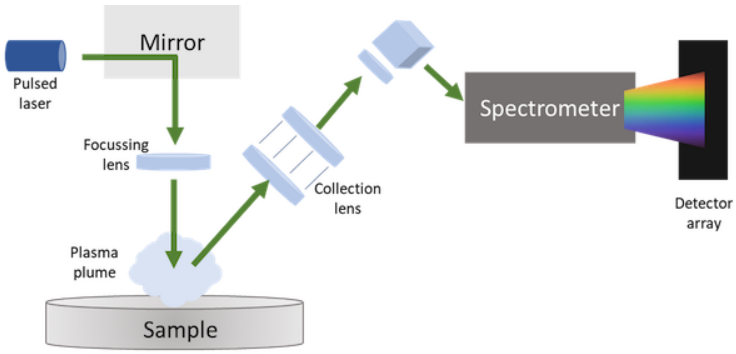

Figure 3: Example of LIBS Setup [3]

A LIBS spectra consists of discrete lines associated with the energy level transitions, specific to the atom or ion. Measurements of intensities of spectral lines from different atoms can be used to determine relative number densities of those respective elements [5].

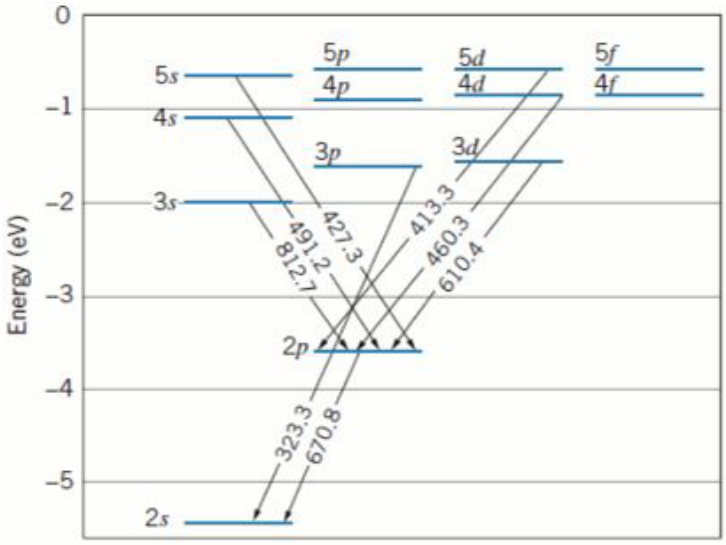

Figure 4: Lithium Energy Level Diagram [4]

The emission lines of interest are lithium, oxygen, hydrogen, and carbon. Others elemental lines may be studied depending on the various parameter that may change within an experiment. For example, if a new buffer gas is used or any contaminants

are known to be in the sample new atomic lines could provide important information. Each of the possible elements in a sample have their own energy level scheme, which determine the emitted spectra. As an example, Figure 4 displays the energy levels for orbital electrons in neutral lithium and the wavelengths associated with these transitions. In the case of LIBS, the laser induces a plasma when the energy of the pulse is imparted on the sample of interest. Such plasma is known as a laser-induced plasa (LIP), and the process is often called laser ablation. As stated before understanding this process in critical to the use of LIBS and the collection of the sample spectra.

2.2.1 Laser Ablation

In addition to forming a LIP, ablation also includes the formation of gaseous vapor and fine particles [5]. LIBS is considered a quasi-nondestructive analysis technique, because during this ablation particles and gaseous vapor are ejected leaving behind a very small crater on the scale of microns. The light emitted from the plasma is the source for LIBS. Understanding the time evolution of this process helps optimization of the technique and involved parameters. The ablation process may be divided into into three major steps which are illustrated in Figure 5.

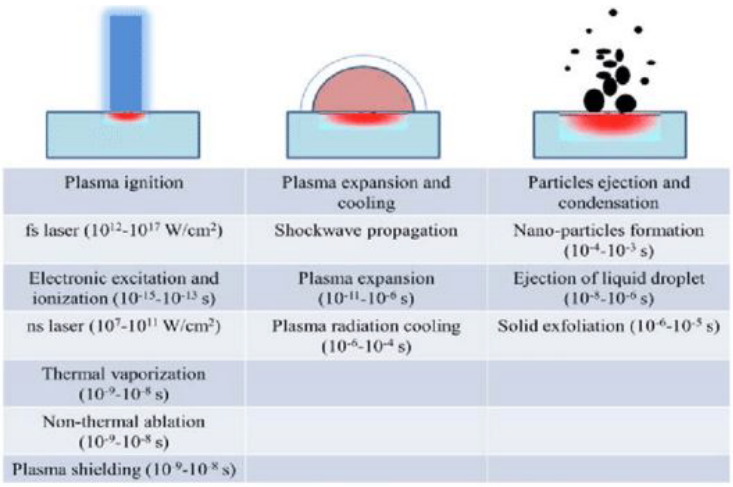

Figure 5: Ablation Process [5]

The first step in laser ablation is the plasma formation, which is done by the deposition of the laser pulse energy into the sample. The primary plasma formation mechanism is thermal for nanosecond lasers [5]. The incident laser causes a rapid rise in surface temperature to the point of vaporization. The vapor continues to absorb energy from the incoming laser pulse until super-heated when ions and free electrons form a plasma. Several factors impact the coupling of energy into the sample. One such factor is the density of the material, while another is the color or the sample material. The more dense the material is, the lower the ablation threshold will be [5]. Also a darker material will induce more energy absorption from the laser pulse.

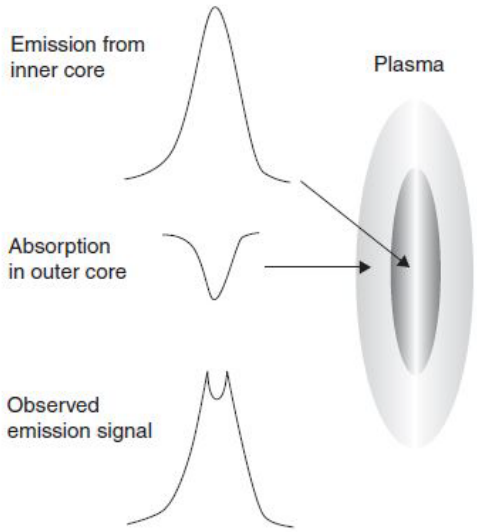

Figure 6: Self Absorption Peak Diagram [6]

The second stage is the plasma expansion and cooling. The plasma continues to expand after the laser pulse for a period of tens of picoseconds to one microsecond following the laser incident on the sample [36]. After the expansion, the plasma begins radiatively cooling. This cooling lasts from microseconds to tens of milliseconds following the laser incident on the sample. During this time frame LIBS spectra is acquired for atomic identification. Prior to approximately one microsecond, the spectra is dominated by a strong continuum background from accelerated electrons and ions, known as the Bremsstrahlung process. The last stage of the ablation process is particle ejection and condensation. This is the step where the ejected materials relax and the crater is formed. These stages and timings set the framework for the experimental setup.

One effect that can also happen during LIBS is when an atomic emission line suffers from self-absorption. This will appear as a flat-top on a peak or, in more extreme cases, a dip known as self-reversal. A visual of this effect on a spectral peak is shown in Figure 6. When the plamsa is not optically thin, the light emitted from the

atoms in the center of the plasma can be absorbed by the cooler atoms of the same element on the peripheral [6]. Wood observed this phenomena in lithium hydroxide in his study of isotopic shifts in lithium [44]. If spectral emission lines have observable self-absorption, there is no longer proportionality between the species concentration and the integrated area under the measured peak. Hence, Atomic emission lines which do not have self-absorption are best when performing analysis. Techniques that reduce self-absorption include performing LIBS under a purge gas such as helium or under vacuum [6, 45]. There are also techniques to handle self-absorption that can be employed during post-processing [6]. Due to sample ablation, the crater does not have the same composition or density as the surface prior to ablation. The thermal propagation into the surrounding material, changes the chemical, molecular, and chromatic properties, which will also change any addition plasma emission [6, 46]. These issues will increase shot-to-shot variation in LIBS and decrease the resolution of the technique for LIBS performed on the same sampling positions. Laser induced breakdown spectroscopy, ever since it was pioneered in 1983 by Cremer and Radziemski at Los Alamos National Laboratory, has been the most extensively used state-of-the-art method of elemental analysis [5].

2.3 Study of Lithium Compounds

As seen in the reaction defined in Eqn 1, when lithium hydride is exposed to moisture it reacts violently and releases gaseous hydrogen. This garners the need for LiH to be handled in a controlled environment due to the release of gaseous hydrogen, which creates a spontaneous combustion hazard [15, 39]. Once lithium hydroxide has been produced by this reaction, the reaction can proceed as shown in Eqn 4, Eqn 5, and Eqn 6. These Eqns show the conversion of $LiOH$ into Li_2CO_3. Due to the reactive nature of lithium compound, this work focuses on moisture and temperature

exposure. The formation of $LiOH \cdot H_2O$ and its conversion to Li_2CO_3 cannot be understated in its importance [14]. Li_2O forms at the boundary layer between LiH and $LiOH$ as shown by Chu [17]. This reaction is described in Eqn 2, and is not represented in the raman spectra due to the Li_2O scattering spectra being outside the range for the optical setup used.

The rates of the various of reactions beginning with LiH is very important for this experiment when determining necessary treatment times. The oxidation of LiH is controlled by nucleation and subsequent growth of $LiOH$. The volume fraction of $LiOH$ as a function of time can then be expressed by the Johnson-Mehl-Avrami (JMA) equation [38]. When considering LiH, the growth of $LiOH$ was obtained by Ren through the curve fitting of as-purchased LiH in air at room temperature as given by the Johnson-Mehl-Avrami (JMA) equation seen in Eqn 7, where f is the growth rate [38].

$$f = 1 - \exp -.18t^{1.01} \tag{7}$$

This places a 90% conversion rate from LiH into $LiOH$ within a few hours. This will be an important parameter when setting environmental exposure times. These conversion rates are influenced by many factors. LiH samples pre-treated with exposure to H_2O, followed by exposure to CO_2 have been shown to have a 50% reduction in rate with any subsequent reactions between LiH and H_2O [29]. LiH samples pre-treated with exposure to CO_2 also have a reduction in the reaction rate, in Eqn 2 [17]. These indicate the formation of Li_2O will likely have less relevance in this experiment due to the overall reduction in this reaction as time progresses. The lithium compounds of interest in this study have been studied with LIBS [21] as well was with Raman spectroscopy [16, 8, 21, 22, 7, 23, 9, 24, 14]. These studies establish the known spectra associated with these techniques and these compounds.

2.3.1 Study of Lithium Compounds using Raman

A wealth of data for each of the molecular compounds of interest in this study have been studied thoroughly over the years using Raman spectroscopy. Raman lines are typically independent of the wavelength of the incident light as long as the incident light is sufficiently different in wavelength than that of the resonance of the excited electronic states (reported in units of cm^{-1}). Certain wavelengths tend to perform better for incident light near resonant wavelengths. It was determined by Stowe that lower wavelengths tend to reduce fluorescence (a common interference in Raman spectra) when considering lithium compounds [20].

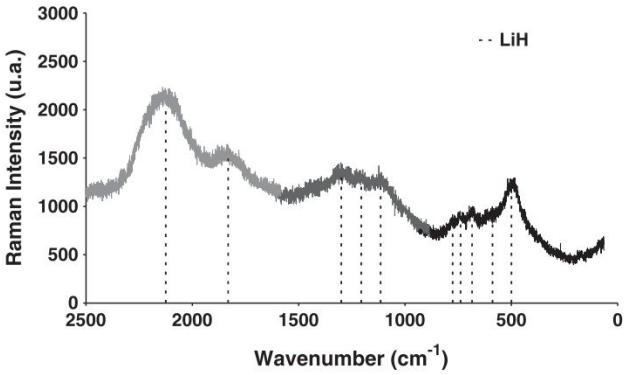

Figure 7: LiH Peaks [7]

A Raman spectra of LiH as determined by Maupoix [7] is displayed in Figure 7. Figure 8 shows an example spectra of $LiOH$ as measured by Gorelik is shown [8]. Figure 9 shows a Raman spectra of Li_2CO_3 as measured by Brooker and Wang [9]. In addition the $LiOH \cdot H_2O$ spectral lines in Table 2.3.1 are crucial in this study.

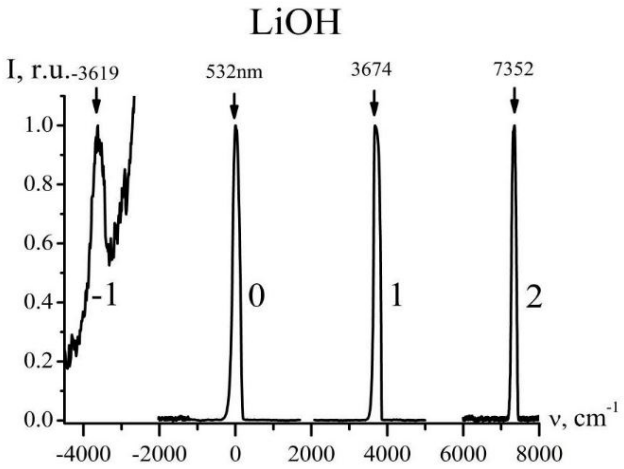

Figure 8: *LiOH* Peaks [8]

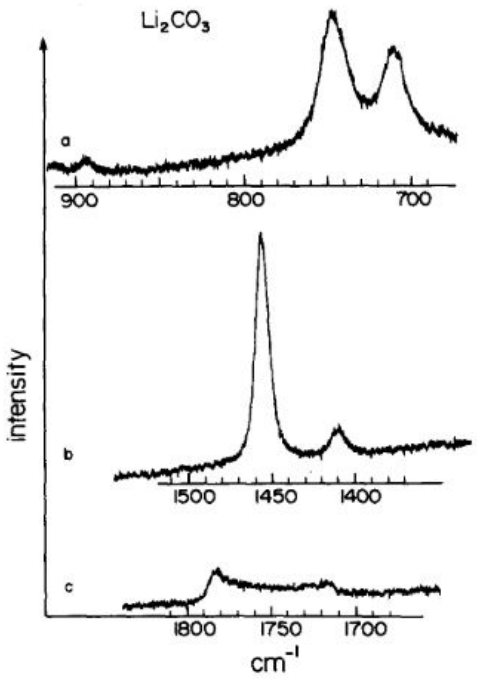

Figure 9: Li_2CO_3 Peaks [9]

A dichroic mirror with cutoff at 550 nm was used to reduce the collection of laser light, which corresponds to 615 cm^{-1}, for 532 nm incident light. The Raman lines of interest in this study will be the Stokes lines, since anti-stokes lines are both weaker

intensity and below the dichroic mirrors cutoff. The optical setups given range, the Raman lines of interest and their corresponding molecular compounds are listed in Table 2.3.1 [8, 23, 9, 14].

Table 1: Raman Spectral Lines of Interest

Raman Spectral Lines		
Compound	λ (cm^{-1})	λ (nm)
LiH	776	554.91
LiH	1228	569.18
LiH	1814	588.82
LiH	2087	598.44
$LiOH$	621	550.18
$LiOH$	3665	660.85
Li_2CO_3	1090	564.74
Li_2CO_3	1461	576.83
$LiOH \cdot H_2O$	863	557.60
$LiOH \cdot H_2O$	3668	656.54

2.3.2 Study of Lithium Compounds using LIBS

Unlike Raman spectroscopy, LIBS is an atomic emission spectroscopy technique. Again, the lithium compounds of interest are LiH, $LiOH$, Li_2O, $LiOH \cdot H_2O$, and Li_2CO_3. Li_2O, while of interest, was not evaluated in this study, due to its Raman spectra being outside the detectable range. The elements of primary focus are lithium, hydrogen, oxygen, carbon, and nitrogen. Nitrogen is of importance due its use as a

fill gas as well as its occurrence in air. Table 2.3.2 displays the elements and the primary atomic lines of interest to which they correspond. Data for selection and identification of these lines was obtained from the NIST LIBS Database [45].

A simulated spectra can be acquired from the NIST LIBS Database, assuming no self-absorption. It can be seen that the carbon line at 833 nm is negligible in comparison to the other carbon lines [47]. Carbon has much stronger emission peaks at 247 nm and 909 nm, but these would require special treatment, due to the fact that these spectral lines are outside the viewed range of the spectrometer in used for this study.

Thus, to overcome the atomic emission lines from carbon not being visible, ratios of the observed peaks can be used to retrieve the chemical information about the lithium compounds present. In a study conducted at Y-12 Security Complex, an experiment was able to use LIBS to observe oxidation of LiH samples by the oxygen atomic line at 777 nm as evidence of oxidation on the surface [21]. Since Li_2CO_3 has both a carbon and oxygen, an increased presence of oxygen and changing ratios of lithium and hydrogen in various mixtures, could similarly be used in the present study to indirectly observe the Li_2CO_3. This is where the value of chemometrics arises.

Table 2: LIBS Spectral Lines of Interest

LIBS Spectral Lines	
Element	Wavelength (nm)
Li	610.354
	610.365
Li	670.776
	670.791
Li	812.623
	812.645
H	656.271
	656.272
	656.277
	656.285
O	777.194
	777.417
	777.539
C	833.514
N	742.364
N	744.229
N	746.831
N	818.487
	818.802
N	821.634
N	824.239

2.3.3 Chemometrics

As used in this study, chemometrics is "the chemical discipline that uses mathematical and statistical methods to provide maximum chemical information by analyzing chemical data" [31]. Using LIBS data, chemometrics has been used in recent years by Stipe and others. To quantify chromium, nickel, and manganese in steel alloys in 2010, Stipe used PLSR and chemometrics [34]. Reviews of chemometrics on LIBS data, were done by Zhang, during the five year window from 2012 to 2016 [33]. Compared to univariate analysis, multivariate chemometrics allows for a more robust predictive models. Univariate methods for creating calibration curves can be fragile due to the susceptibility to fluctuations in laser energy and matrix effects in the sample [33]. Also, chemometrics employed with data fusion, to further enhance the predictive capabilities, have been seen in recent work [48, 49, 50]. For example, NASA is using chemometrics and data fusion on the Mars rover and other solar system exploration modules [49]. These studies provide support for the pursuit of these methods to be employed due to their popularity and accepted use in LIBS and Raman communities.

2.3.4 Data Fusion

Due to the ability to conduct both LIBS and Raman measurements in the same experimental setup, data fusion is an appealing data manipulation technique. Data fusion is when multiple data types are used together to create a more accurate model [51]. There are several categories of data fusion based on the level of data manipulation before fusion occurs [51]. This research uses low-level data fusion where the raw data from Raman and LIBS are fused together prior to any manipulation or characteristic extraction. LIBS and Raman spectroscopy are easily used together due that they are both spectroscopic measurements and they provide complimentary informa-

tion about a sample. That LIBS provides elemental composition information and Raman provides molecular identification. As mention in section 2.3.3, data fusion of LIBS and Raman spectra is seeing increased popularity as demonstrated in its application to terrestrial and martian environmental samples [48, 49, 50]. To identify the presence of minerals and other naturally occurring elements of interest, each of the cited works paired LIBS-Raman data fusion with chemometrics techniques. Similarly, in this study, data fusion and chemometric techniques will be used in combination with various predictive models in this experiment.

2.3.5 PCA and PCR

PCA is a method to dimensionally reduce a data set. This is done by finding or creating latent variables which explain the bulk of the variance in a data set. The advantage to using principle components for regression analysis is they allow for the reduction of multicollinearity [52]. Principal components regression takes the output of a PCA and performs regression analysis on those principal components [53]. When multiple explanatory variables correlate with each other it is referred to as multicollinearity [54]. This is important because multicollinearity is known to lead to unstable and unreliable estimates of regression coefficients. By selecting only the principle components which explain the bulk of the variance, multicollinearity can be minimized, allow for increased modeling power [54].

When choosing principle components, first one can decide how much variance must be explained by a PC in order to be included. A principle component might not be included in the regression model if it fails to explain at least the determined level of the total variance. Secondly, one can decide how much total variance must be explained (e.g. 90%). In this case, the PCs which explain the highest portion of the variance are included until until 90% of the variance is explained. Once the principle

components are chosen, a linear least squares fit is performed, createing a regression model for the data set. LIBS, Raman data, and data fusion have been used with PCR separately and in combination with eachother [44, 48, 55].

2.3.6 PLSR

While PCR is a method for regression modeling. It has the drawback that the dependent variable (in the case of this study it is the environmental traits of the treatments) is not considered when constructing the principle components. This create the possibility that the principle components will be able to accurately account for the variance in the data set, but will not accurately predict the response variables. Unsupervised techniques such as PCR typically have this disadvantage [56]. Hence the benefits available in supervised learning techniques such as PLSR.

PLSR is a supervised learning technique, which like PCR also reduces the dimensionality of the data by finding latent variables based the raw data set of variables. It identifies and builds a new set of features which are linear combinations of the original variables. The main distinction between PCR and PLSR is that the selection and weighting of these new variables and feature space are governed by the response variable [56]. This is accomplished by minimizing the covariance between the predictors and the responses [57].

While PLSR has become very popular in chemometrics, it also has a trade-off that occurs [44, 48, 49, 50, 55]. As a supervised technique, PLSR reduces bias amongst the variables, but as a predictive model it can suffer from over-fitting due to the inclusion of response variables in model construction. These trade-offs versus PCR can reduce, if not completely negate the advantages of PLSR, suggesting one technique is not better but simply has different strengths and weaknesses [56]. This will dictate a comparison between the techniques employed to judge how each handles the data set

presented in this experiment.

2.3.7 Random Forest

Decision trees are another modeling technique for multivaiate analysis. Random Forest bags an ensemble of decision tree to be used in either classification or regression modeling. Bagging stands for bootstrap aggregation, which is where each independent tree in the ensemble is grown from a bootstrapped replica of input data. Observations not included in this replica are "out of bag" (OOB) for this tree, providing validation within the model construction. Each tree is built using a random forest decision tree process, which is a nonparametric multivariate technique based on an machine-learning algorithm [58].

In comparison to PLSR, the Random Forest has several advantages which make it an appealing predictive model with the data sets in this experiment, which are listed below [58]:

- It allows the employment of both categorical and numerical variables.
- No assumption is required about the statistical distribution of the data.
- It is capable of accounting for interactions and nonlinearities between variables.
- It can be considered robust with respect to changes in the composition of the dataset.
- Prediction with OOB data avoids overfitting.
- It allows exploring a large number of explanatory variables because it intrinsically emphasizes only those variables of high explanatory power at each node split.
- It is robust with respect to noise features.

The use of OOB data to build each decision tree, allow for less vulnerability to over fitting, which is an area where PLSR has difficulty, making the comparison between Random Forest and PLSR of particular interest. Random Forest provides the ability to use categorical response variables which allows Raman and LIBS data to be modeled for classifications using a combination of response variables (e.g. humidity, temperature, and fill gas).

III. Methodology

This experiments methodology was broken down into the four sequential stages used: sample preparation and environmental treatments, Raman setup, Laser Induced Breakdown Spectroscopy (LIBS) setup, and data processing.

3.1 Environmental Treatments and Sample Preparation

For this study of LiH, a PLAS-LABS MODEL 890-THC/DT Glove Box was employed. This environmentally controlled glove box was used to control the humidity and temperature for sample exposure and preparation.

Figure 10: Environmental Chamber

The LiH used in this study was acquired from Sigma-Aldrich in powdered form. The powder form was used due to it being readily available from previous experimental work done, as well as the ability to customize sample preparation for the treatment of the sample materials.

Due to this experiments focus on LiH and its exposure temperature and humidity, liquid nitrogen boil-off was used to purge the environmental chamber, shown in Figure 10. This was done to create a dry, none reactive environment as a beginning point for this study [22, 39]. While purity of the nitrogen boil-off is unknown, periodic purges were done to through the duration of the sample treatment and preparation to keep the O_2 levels below 3%. The O_2 was measured by a PureAire TX-1100-DRA Oxygen Deficiency Monitor. By measuring the O_2 levels, a relative level of CO_2 could be estimated. This was a concern due to the reaction between $LiOH$ and CO_2, which was initially trying to be controlled. Through minimization of the presence of other constituents, the humidity and temperature effects on the LiH samples can be isolated based on the known reactions of LiH [39]. In a later stage of the experiment these nitrogen purges were eliminated to allow for CO_2 exposure. While the use of other inert gases such and helium and argon have been shown to improve LIBS measurements the [21, 59, 60], nitrogen was used based on the fact that it is the most common purge gas for LiH handling and storage, and it is also the most widely available purge gas on the market.

The equipment used in the glove box consisted of a lab scale with 1 mg precision, a mortar and pestle, a powder sample mixer, and a pellet press. Due to the space constraints of the glove box, each piece of equipment was chosen based on their size to allow for sample preparation with the confines of the environmental chamber. This impact on the equipment size affected the precision of the scale, the size of the sample pellets, and the maximum pressure of the press. However, the scale's precision was acceptable because the mass of each lithium compound in the samples was at least 100x the precision of the scale so as to reduce the percent uncertainty in the measurement below 1%.

3.1.1 Environmental Treatments

This experiment entailed using 25 different treatments of LiH to study the effects of temperature, humidity, atmospheric composition, and exposure time. These parameters are shown in Table 3.1.1.

Table 3: LiH Environmental Treatment Parameters

Treatments Parameters					
Label	Humidity (RH%)	Temp. ($^\circ F$)	O_2(%)	Fill Gas	Exposure Time
TM0	<5	65	<3	Nitrogen	0hr
TM1	<5	65	<3	Nitrogen	2hr
TM2	<5	65	<3	Nitrogen	24hr
TM3	<5	85	<3	Nitrogen	2hr
TM4	<5	85	<3	Nitrogen	24hr
TM5	<5	105	<3	Nitrogen	2hr
TM6	<5	105	<3	Nitrogen	24hr
TM7	35	65	<3	Nitrogen	2hr
TM8	35	65	<3	Nitrogen	24hr
TM9	35	85	<3	Nitrogen	2hr
TM10	35	85	<3	Nitrogen	24hr
TM11	35	105	<3	Nitrogen	2hr
TM12	35	105	<3	Nitrogen	24hr
TM13	70	65	<3	Nitrogen	2hr
TM14	70	65	<3	Nitrogen	24hr
TM15	70	85	<3	Nitrogen	2hr
TM16	70	85	<3	Nitrogen	24hr
TM17	70	105	<3	Nitrogen	2hr
TM18	70	105	<3	Nitrogen	24hr
TM19	<5	85	20	Air	2hr
TM20	<5	85	20	Air	24hr
TM21	35	85	20	Air	2hr
TM22	35	85	20	Air	24hr
TM23	70	85	20	Air	2hr
TM24	70	85	20	Air	24hr

Each treatment begins with weighing of two separate LiH samples of .250g $\pm .01g$. This is done after the environmental chamber is set to the proper settings and is in a stable state, humidity's $\pm 1\% Rh$, and $\pm 1^\circ F$. After 2 hours the first sample group enters the pellet pressing process and then data collection, while the second sample

group continues exposure for a full 24 hour exposure, at which point it enters the pellet making process and then data collection.

3.1.2 Pellet Pressing Procedure

After environmental exposure, the treated powders were ground with a mortar and pestle prior to pressing to ensure uniform particle size [61]. Tawfik demonstrated the utility of grinding powder samples prior to pressing in their study on optimization of pellet pressing parameters for use in LIBS [47]. After grinding, the samples were mixed in a Fluxana MUK mixer to ensure homogeneity. Following mixing, the powder was transferred in small quantities into the press die. The die is 7 mm in diameter and produces, on average, a pellet 1.5 mm in thickness. The pellet press is a Specac mini pellet press with maximum pressure of 2 tons. The pellets were pressed to 1.5 tons and this pressure was sustained for 3 min. For the 3 min period pressure was continually adjusted to maintained 1.5 tons, which allowed for pellet settling until the pressure was holding constant.

Figure 11: Sample Containment Vessel

After samples were pressed, they were moved to the conflat flange cell using a thin layer of vacuum grease to adhere the sample to the cell. This grease kept the samples from shifting in the cell while transferring to data collection. Also during LIBS, the grease kept the samples from shifting due to the violent ablation process.

The samples were placed in this air-tight cell with a view port in order to allow

analysis of the samples without exposure to a reactive environment. The view port allowed the laser to interact with the sample as well as for light collection from the samples. The Conflat flange sample cell has a UV fused silica view port with spacer to extend the distance between the focal point of the laser being focus on the sample, creating more distance from the view port to sample in order to reduce the risk of interaction with the view port. By performing data collection immediately following sample preparation the potential impacts to the sample integrity were minimized.

3.2 Dual Raman and LIBS Setup

For each treatment, 4 samples were made and each sample was shot in 5 different areas for data collection, using Raman Spectroscopy and LIBS. This is shown in Figure 12.

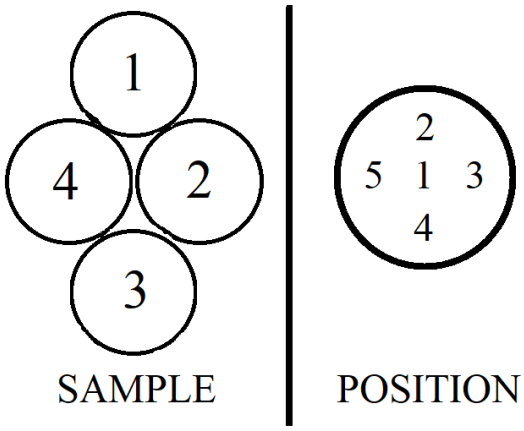

Figure 12: Sample and Shot Positions Diagram

Due to each individual sample being used for both Raman and LIBS measurements, Raman measurements needed to be obtained prior to any LIBS measurements. This is due the destructive nature of the ablation process in LIBS, which would effect any Raman measurement taken afterward. Using the combined optical setup shown in Figures 13 and 14 along with an Echelle Spectrograph, allowed for Ramman and

LIBS sprectra to be collected with minimal setup alterations.

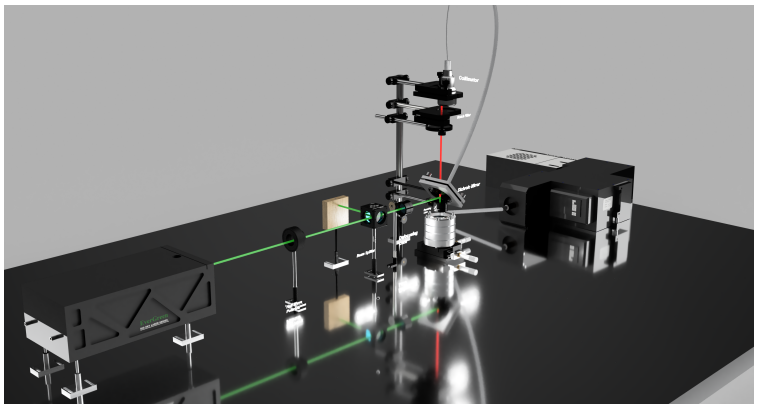

Figure 13: Combined Raman and LIBS Setup

A beam polarizer and beam splitter were used to control laser intensity to evaluate setting parameters during setup optimization. A 550 nm high-pass dichroic mirror was employed in conjunction with a series of notch filters to reduce the collection of laser light as discussed in subsection 2.3.1. While these were needed specifically for the Raman measurements they were not adjusted during LIBS measurments. The only adjustment needed between Raman and LIBS measurments was the removal or addition of a defocusing lens and a change in the laser power.

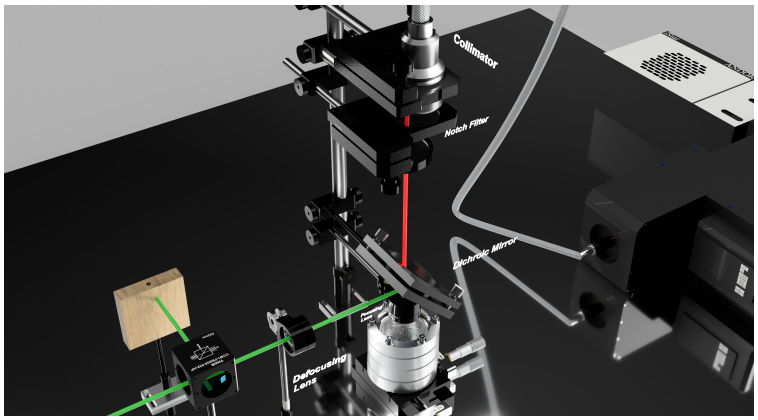

Figure 14: Combined Raman and LIBS Setup Close Up

3.2.1 Echelle Spectrograph for Raman and LIBS

In comparison to a standard monochromater, an echelle spectrograph uses 2-D diffraction rather than 1-D diffraction and it has a much broader bandwidth range. For viewing multiple lines in a broad bandwith a Czerny-Turner style monochromater, it would require adjusting the grating between multiple images on the camera to allow for observation of two distinct a spectra. This requires multiple exposures and in LIBS this means separate ablations which is not viable due to the damage caused by ablations. With an echelle spectrograph, the bandwidth is hundreds of nanometers, which gives greater bandwidth viewing ability versus the monochromater. The echelle spectrograph used in this study is a Catalina Scientific EMU-120/65 model, with a range of 325 nm to 925 nm. In chemometrics with LIBS data, this broad bandwidth is especially useful due to the number of elements of interest in a single spectra. At least some spectral lines from nearly every element of interest in the LiH reactions were present in this range making it possible to collect data vital information from these atomic spectral lines.

This broad spectral range also benefits the use of Raman spectroscopy. Since Raman is based on a wavelength shift from an excitation laser wavelength, as long as the excitation laser is at the lower end of the spectrographs range, then the Stokes shift should be visible in the spectra. For example, a standard 1064nm laser would not be a valid excitation for this setup due to it being outside the upper limit of this spectrograph. For this study, the largest Raman shift is approximately 3600 cm ($LiOH$), with an excitation laser wavelength of 532 nm the shift of 3600 cm put a spectral line around 658 nm. This is well within the spectral range of the echelle spectrograph which makes the selection of a 532 nm excitation laser appropriate when looking to utilize Raman spectroscopy. This combines well with the higher energy of a 532 nm laser, which provides the needed fluence when LIBS is being used as a measurement

technique.

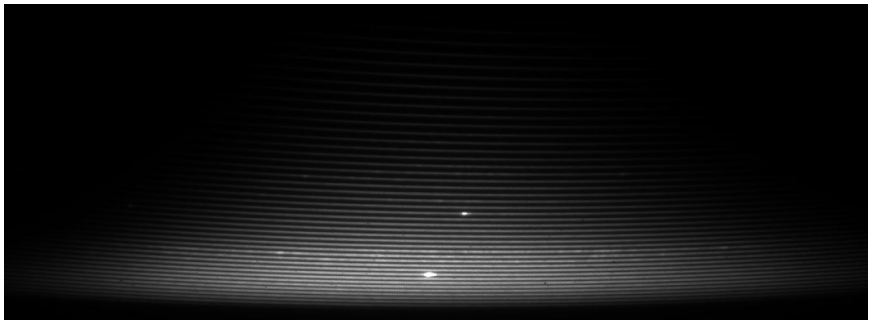

Figure 15: Deuterium and Tungsten Halogen Spectrum Used for Flat-Field Correction

When using the Catalina Scientific EMU-120/65 model echelle spectrograph, KestrelSpec data acquisition and analysis software also developed by Catalina Scientific, was used. Calibrations were performed before each data collection, which corresponds to one calibration per 5 hours of operation. First, a wavelength calibration was performed using a Hg/Ar calibration lamp. Along with the wavelength calibration, an order optimization and flat-field calibration was also performed. The order optimization uses an image of a continuum light source (deuterium tungsten halogen) to inform the optimization of the cutoff between orders using an internal algorithm in KestrelSpec. An image of a deuterium/tungsten and halogen lamp is displayed in Figure 15. In addition to order optimization, a flat-field calibration is performed using the tungsten halogen lamp to approximate a black body radiator.

3.2.2 Raman

The Raman spectra were acquired via integrating images in a single exposure and then accumulating multiple exposures to form the spectral image. This was done at a repetition rate of 10 Hz for 600 pulses per exposure with 5 exposures. The gate delay and gate width of the ICCD camera were optimized for capturing the scattered

light. The exposure time, which is how long the camera shutter is open, is set to its minimum value (1 ms). Microchannel plate (MCP) gain is adjustable from 0-4000 on a non-linear scale. Raman is a scattering process which is why the timing of collection is on the nanosecond scale. Laser light collected with these reflections is reduced by a 550 nm high-pass dichroic mirror and notch filters in the optical setup. This reduces the saturation by the laser light which creates a bigger signal to noise ratio between the Raman spectra and the excitation laser. The settings used for the Raman measurements taken can be found in Table 3.2.2.

Table 4: Raman Setup Parameters

Raman Settings			
Equipment	Parameter	Value	Sync
Laser	Power	100.4mJ	NA
	Flashlamp Trigger	0 μs	T0
	Q-Switch Trigger	190μs	T0
ICCD	Camera Trigger	0 s	Q-Switch Trigger
	Exposure Time	60s	Camera Trigger
	Integrated Pules	600	N/A
	Gate Delay	100ns	Camera Trigger
	Gate Width	40ns	Camera Trigger
	MCP Gain	2500	N/A

3.2.3 LIBS

Following Raman measurements, LIBS measurements were taken. An unperturbed position was chosen for the beginning of each set of shots. Each set of shots consisted of 5 ablations in a single position. Each ablation was recorded individually and not integrated or accumulated on camera like the Raman spectra. Each sample was ablated in 5 separate positions for a total of 20 spectra per treatment. The timing of LIBS measurements differs in comparison to those associated Raman. Where Raman is a scattering process with nearly zero delay needing a collection on the scale

of nanoseconds after excitation, the mechanisms of ablation, plasma formation, and relaxation which generate the LIBS spectra take microseconds to occur as discussed in Section 2.2.1. The LIBS parameters for this study can be found in Table 3.2.3.

Table 5: LIBS Setup Parameters

LIBS Settings			
Equipment	Parameter	Value	Sync
Laser	Power	151.9mJ	NA
	Flashlamp Trigger	0 μs	T0
	Q-Switch Trigger	190μs	T0
ICCD	Camera Trigger	0 s	Q-Switch Trigger
	Exposure Time	.001s	Camera Trigger
	Gate Delay	1μs	Camera Trigger
	Gate Width	6μs	Camera Trigger
	MCP Gain	2000	N/A

3.3 Analysis Techniques

In this experiment data manipulation and analysis were performed in Matlab version R2021b.

3.3.1 Raman and LIBS Data Fusion

Each treatment includes 4 samples with 5 shots per sample, or 20 data points per treatment for both Raman and LIBS, producing 500 data sets in total. For each significant peak within the different spectra a 5 nm window is taken around the centroid of the peak. If multiple peaks are grouped close together a 2.5 nm range is used on either side of the outer most peak for a the collection of peaks. The wavelength step per channel was .01 nm. This produces data sets of 500 samples environmentally treated, with a corresponding 4381 spectral channels, with .01 nm width, for the Raman spectra and 4529 channels for the LIBS spectra before data

fusion. Prior to fusion each spectra was normalized by the max intensity of the spectra in order to prevent features of one spectra from becoming overly weighted during analysis. Analysis was performed on each data set to provide a comparison the different data types and how well the model the lithium conditioning.

After analysis of Raman and LIBS data, data fusion was completed on the data obtained from LIBS and Raman measurements. For the data fusion each Raman spectra was appended to the end of the corresponding LIBS spectra, second shot since the first shot cleans the sample, from the same exact pellet and same position. This created pseudo-spectra data set representing each sample and position for a 500x8910 data set. With this data fusion of the respective Raman and LIBS spectra, the multivariate analysis techniques were performed as were used on the original data sets.

3.3.2 PCA

Principle components analysis (PCA) was also performed to to establish how well the different traits of the treatments appear in principal component space. If separation is not observed, the likelihood of an accurate regression model using PCR could be less than optimal. PCA was also used to present a visual representation of different traits seen in Raman and LIBS spectra, and fused data to possibly see trends in the data.

3.3.3 PCR and PLSR

In multivariate analysis, PCR and PLSR were conducted using humidity, temperature as response variables. These used 25 components when modeling, which was determined by a investigation into the component space for each model. PCR and PLSR models were built using Raman, LIBS, and LIBS-Raman fused data for

comparison. Multiple multivariate models were built with humidity and temperature data to explore the best method for the data collected. The metric for effectiveness of the models was both the R^2 value and the Root Mean Square Error (RMSE) of each linear fit. The R^2 value is a measure of goodness of fit between a regression line and the data it is being fit to and is given by Eqn 8 where RSS is the Residual Sum of Squares and TSS is Total Sum of Squares. The smaller the residuals, the closer the R^2 value is to 1.

$$R^2 = 1 - \frac{RSS}{TSS} \tag{8}$$

The RMSE is a representation of the difference between the observed and the predicted values. More specifically, RMSE is the standard deviation of the residuals of a model and can be described by Eqn 9. Better performance of the model is indicated by higher R^2 and lower RMSE.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}} \tag{9}$$

3.3.4 Random Forest

Lastly, a decision tree method known a Random Forest was done characterizing each response variable into a categorical format for humidity, temperature, treatment time, and atmospheric composition. In addition, a Random Forest regression was done for humidity and temperature responses. Random Forest bags an ensemble of a dictated number of decision trees for either classification or regression. As every tree in the ensemble is grown not all observations are included in the training. Any data not used to train a tree is considered "out of bag" for this tree. Random Forest relies on the ClassificationTree and RegressionTree functionality for growing individual trees in Matlab. In particular, ClassificationTree and RegressionTree accepts the number

of features selected at random for each decision split as an optional input argument. This means that each decision tree is built with a random feature picking order.

By default all the input data is used to sample with replacement when growing each new tree, which was altered to use a 80% in bag and 20% out of bag split. The number of variables to selected at random for each decision split is the square root of the number of variables for classification and one third of the number of variables for regression. The minimum number of observations per tree leaf is 1 for classification and 5 for regression. For the regression models a linear fit was done and the R^2 and RMSE were found to provide a comparison to the PCR and PLSR models used previously.

This model was initially implemented with no training-test split to be comparable to the PCR and PLSR methods used previously. In addition to this a training-test split was performed using 80% of the initial data for Raman, LIBS, and the Raman-LIBS fusion data for training, with 20% held for model testing.

IV. Results and Analysis

This section will focuses of the multivariate analysis of the Raman, LIBS, and LIBS-Raman Fusion, with the response variables being humidity, temperature, exposure time, and atmospheric type.

First, the spectral responses in the Raman and LIBS spectra are viewed to breakdown the different trends that will hopefully be seen in later modeling, as well as any unexpected spectra that may occur. PCA is then used to see the separation of these trends in a principle component space and to provide a initial visualization of how well PCR and PLSR will perform on the individual data sets. PCR and PLSR are used to provide regression modeling for humidity and temperature due to the continuous response. Training and test splits used an 80/20% split, and was performed to provide a more realistic data set for testing.

Next, a Random Forest model is employed as an alternate modeling technique which can use both categorical and continuous response variables. This was done to provide an additional regression model as a comparison to PLSR, but also to provide an example of the general classifications that could be performed based on a varying response variables and grouping of response variables. When looking at PCR, PLSR, and Random Forest Regressions, R^2 is used as a relative measure of goodness of fit, while $RMSE$ is an absolute measure of fit, RMSE is viewed as the standard deviation of the unexplained variance. Similar to the PCR and PLSR training-test split, a mew test data set was applied to models constructed.

4.1 Raman Data Analysis

Analysis was performed using only the normalized Raman spectra using spectral analysis, PCA, PCR and PLSR, and the Random Forest modeling techniques.

4.1.1 Raman Spectra

Using the main lines in Table 2.3.1, the Raman Spectra is investigated for visual trends during the environmental treatments. Higher temperature and humidity settings are viewed due to the more exaggerated changes between each spectra.

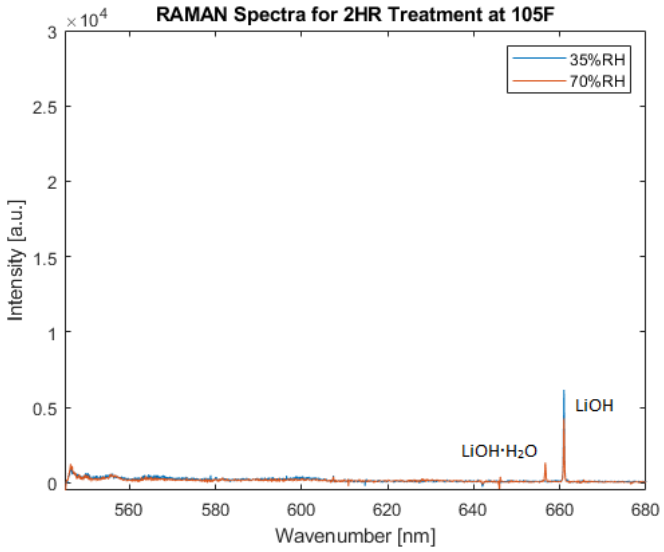

Figure 16: Raman Spectra for 2hr Treatment

When viewing the 2 hour treatments at 35% and 70% humdity in Figure 16, as the humidity was increased the $LiOH$ peak at 660.85 nm initially increased, but eventually began to decrease and then the peak for the $LiOH, H_2O$ peak at 656.54 nm can begin to be seen. As, the treatment time is increase this trend continues with the addition of the $LiOH, H_2O$ line at 557.60 nm also becoming visible in Figure 17.

As the nitrogen purges were discontinued in later treatments, there is a clear Li_2CO_3 peak at 564.74 nm, and the $LiOH, H_2O$ at 656.54 nm decreases and the $LiOH$ peak at 660.85 nm increases as seen in Figure 18. The $LiOH$ peak increase is thought to be due to the lower temperature and humidity which slows the conversion of $LiOH$ to $LiOH_2O$ as well as that the conversion of LiH to $LiOH$ could still be occurring for the lower humidity of 35%. The $LiOH \cdot H_2O$ and $LiOH$ both

should decrease as the reaction to Li_2CO_3 occurs, this is only clearly visible in the $LiOH, H_2O$ peak when comparing Figure 17 and Figure 18.

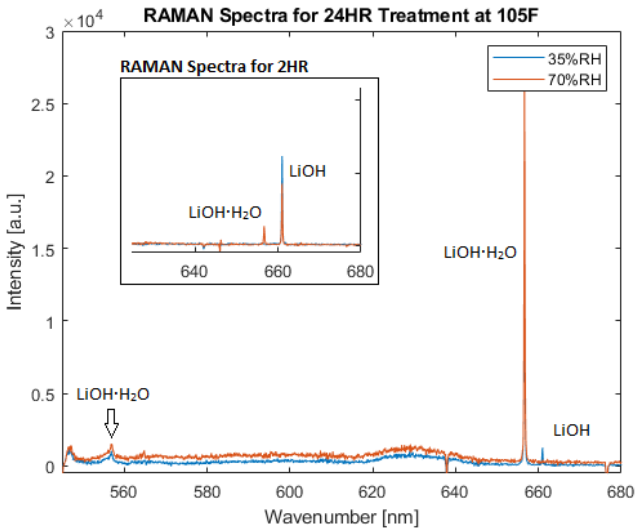

Figure 17: Raman Spectra for 24hr Treatment

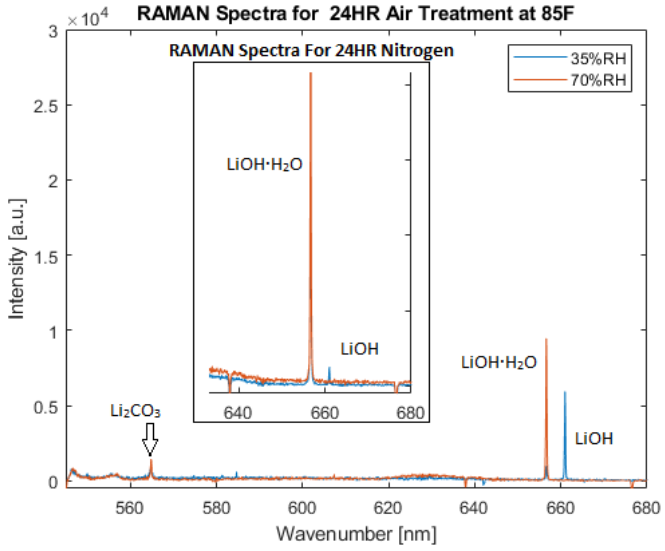

Figure 18: Raman Spectra for 24hr Treatment in Air

4.1.2 Raman Data PCA

Using the Raman spectral data, PCA is used to provide a gauge of the separation in a component space and the viability of modeling using this data. This is done while keeping the response variables of overall treatment classification, humidity, temperature, exposure time, and atmospheric type in mind. Treatment settings can be referenced in Table 3.1.1, for clarity of these separations.

In Figure 19 there is a clear separation between in the different treatments. For example, treatments 8,10,12 (in green) are 35% humidity treatments, with nitrogen atmosphere, for 24 hour exposure, and a clear separation is seen from treatments 2,4,6 (shown in blue) which only have an increase in only humidity between them. As seen in Figure 20, solely humidity does not explain the separation in all cases.

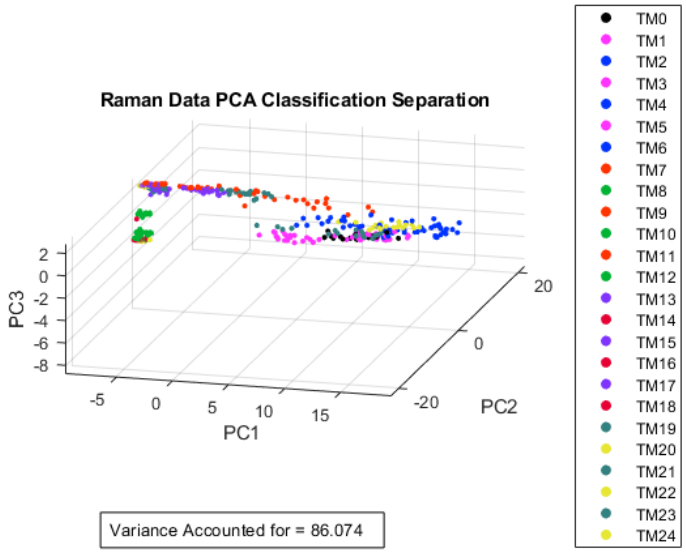

Figure 19: Raman PCA for Classification Separation

Low humidity is clearly separated (in blue) but the 35% (in green) and 70% (in red) humidity data is somewhat mixed. However, by also using Figure 21, the exposure time combined with humidity can explain much of the variance in the PCA plots.

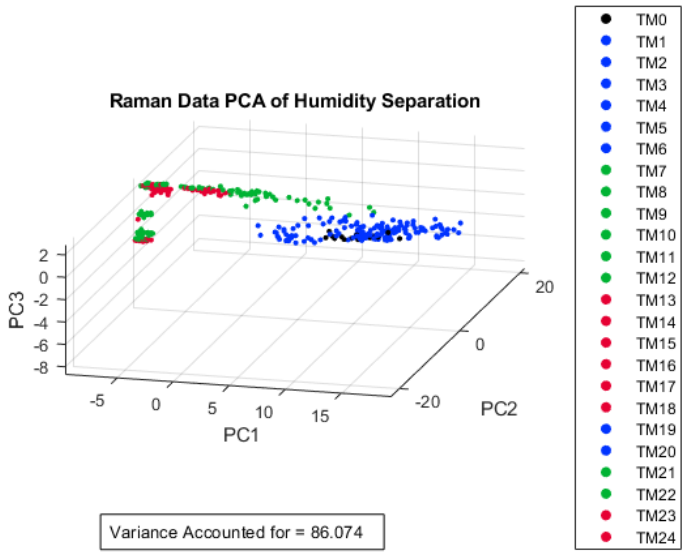

Figure 20: Raman PCA for Humidity Separation

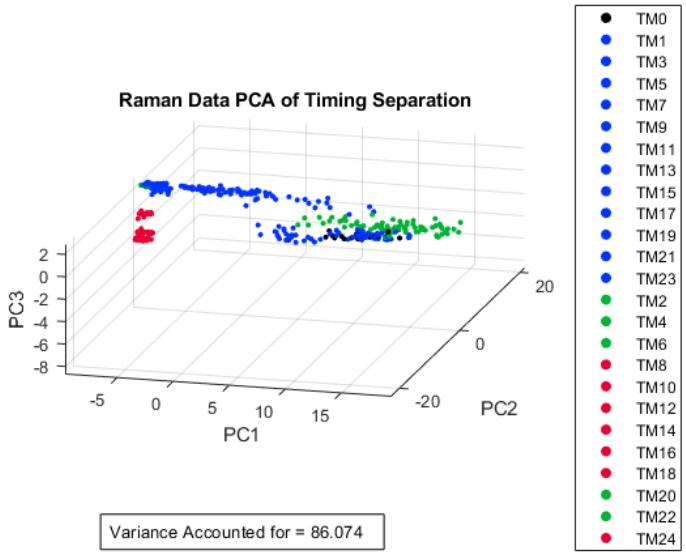

Figure 21: Raman PCA for Exposure Timing Separation

Temperature did not create clear separation as a response in PCA due to it being more of a secondary effect, with humidity and exposure separation being more

distinguishable. This seems to indicate that there will be adequate separation in the data to provide significant modeling power.

4.1.3 Raman Data PCR and PLSR

Using PCR and PLSR modeling, the Raman spectral data was applied to humidity and temperature as the only two continuous responses focused on in this experiment.

4.1.3.1 Raman Data PCR and PLSR for Humidity

More clarity can be distinguished about how well PCR and PLSR will perform with the Raman data, by viewing the first two components after feature reduction by each model. This is paired with the response variable in question to provide insight.

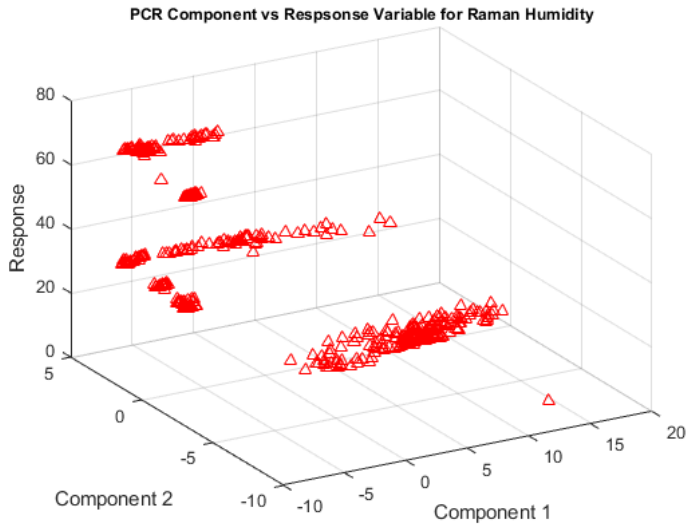

Figure 22: Raman Data PCR Component Space Versus Humidity

When comparing the first two components of PCR and PLSR with reference to humidity, one can see that their is comparable separation in both component spaces as through Figure 22 and Figure 23.

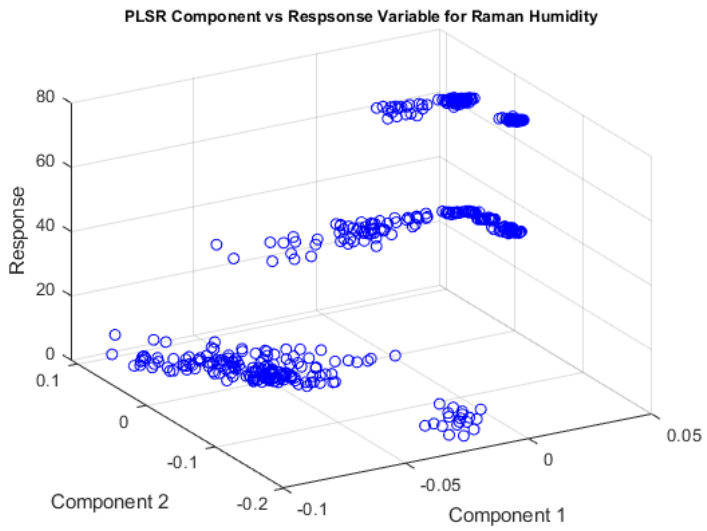

Figure 23: Raman Data PLSR Component Space Versus Humidity

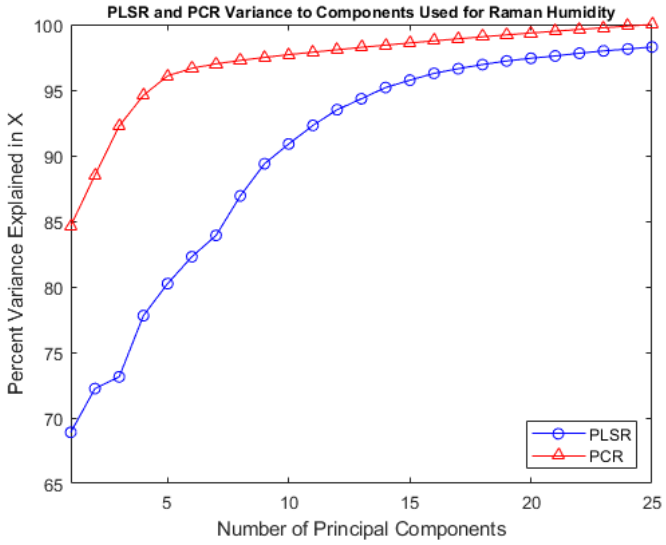

Figure 24: Raman Data PCR and PLSR Variance Per Component for Humidity

By using Figure 24 an appropriate number of components was picked to use when modeling for PCR and PLSR accounting for approximately 95% of the variance.

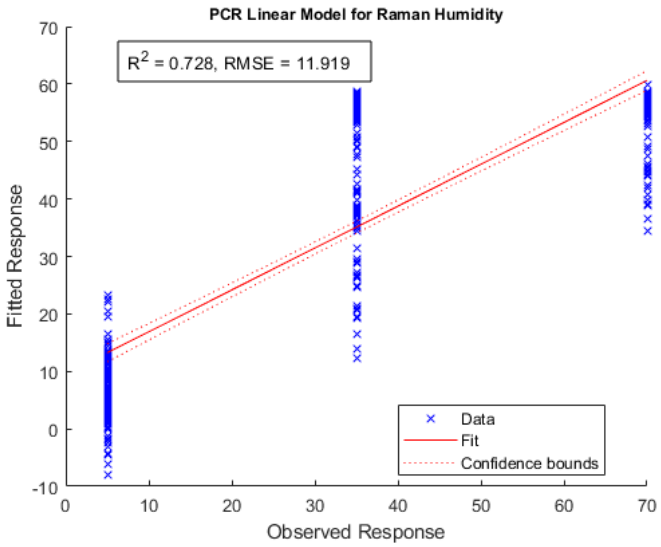

Figure 25: Raman Data PCR for Humidity

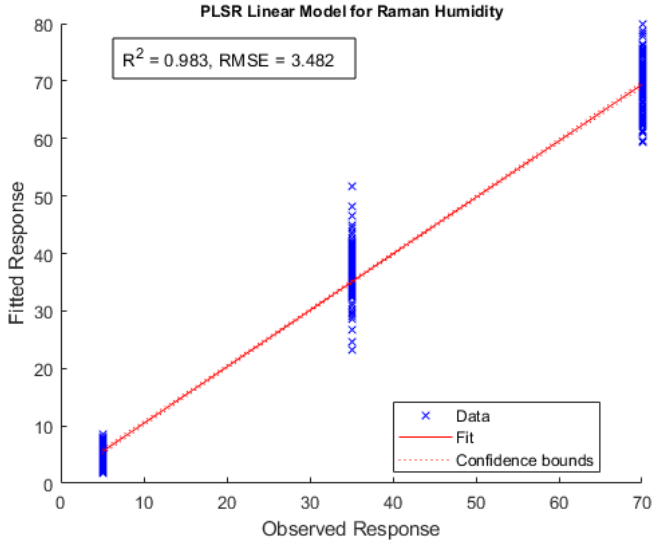

Figure 26: Raman Data PLSR for Humidity

In general, PCR performs slightly worse than PLSR. PCR had an $R^2 = .728$ and a $RMSE = 11.919$, while PLSR had an $R^2 = .983$ and a $RMSE = 3.482$, as seen in Figure 25 and Figure 26. This make some sense in by the fact that PLSR has less significance per component on average, meaning it can more accurately describe the

variance in the data, without over simplify the data separation.

4.1.3.2 Raman Data PCR and PLSR for Temperature

When comparing the first two components of PCR and PLSR with reference to temperature, the PLSR separation in component spaces seem much cleaner. This is shown by a comparison of Figure 27 and Figure 28.

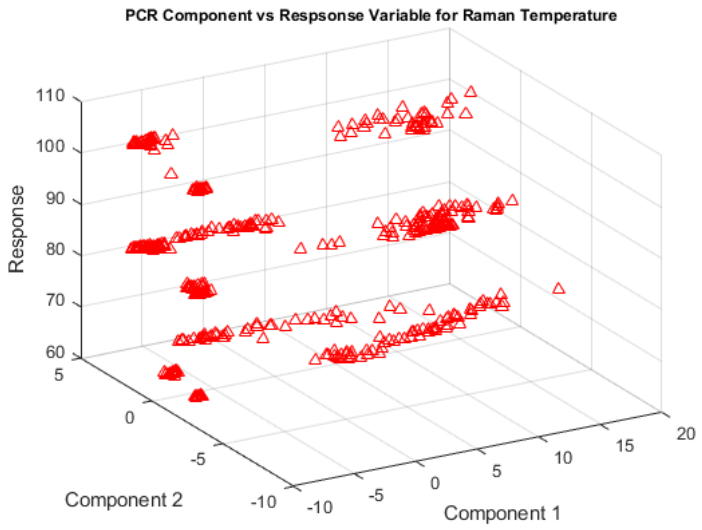

Figure 27: Raman Data PCR Component Space Versus Temperature

At the same time in Figure 29, PCR uses only approximalty 4 components before accounting for almost 95% of the variance, while PLSR needs 25 components to reach the same amount of variance explained. As explained in the section 4.1.3.1, this also leds to better performance when using PLSR for a temperature response.

for the PCR and PLSR models, using R^2 and RMSE as the measure of goodness and validity of the model, PCR perform significantly worse than PLSR. PCR had an $R^2 = .454$ and a $RMSE = 7.186$, while PLSR had an $R^2 = .950$ and a $RMSE = 3.138$, as seen in Figure 30 and Figure 31

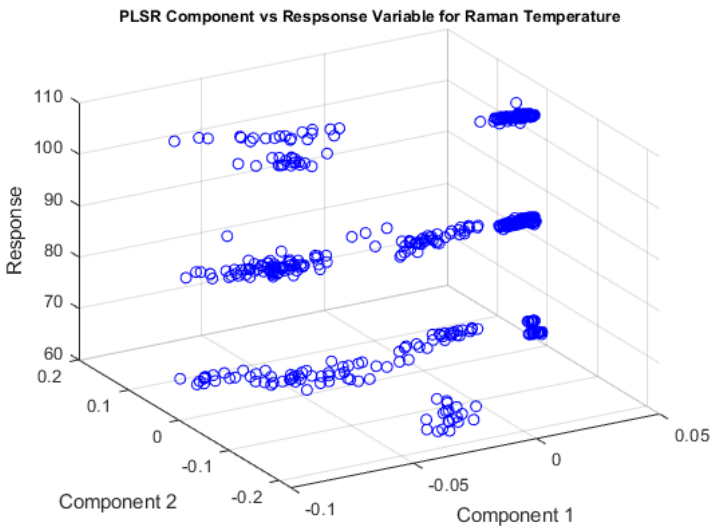

Figure 28: Raman Data PLSR Component Space Versus Temperature

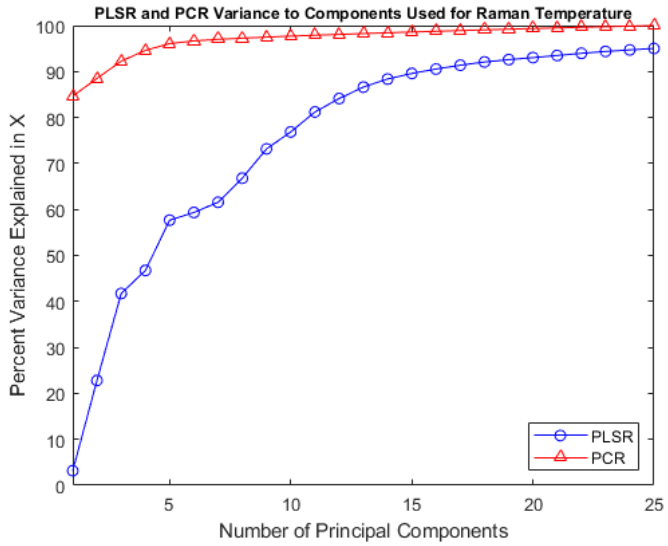

Figure 29: Raman Data PCR and PLSR Variance Per Component for Temperature

When comparing the PCR and PLSR humidity models to the PCR and PLSR temperature models, the humidity models perform better. Especially the PCR model for humidity with an $R^2 = .728$ over temperature model with an $R^2 = .454$. This is due to the clear difference in separation in the component space.

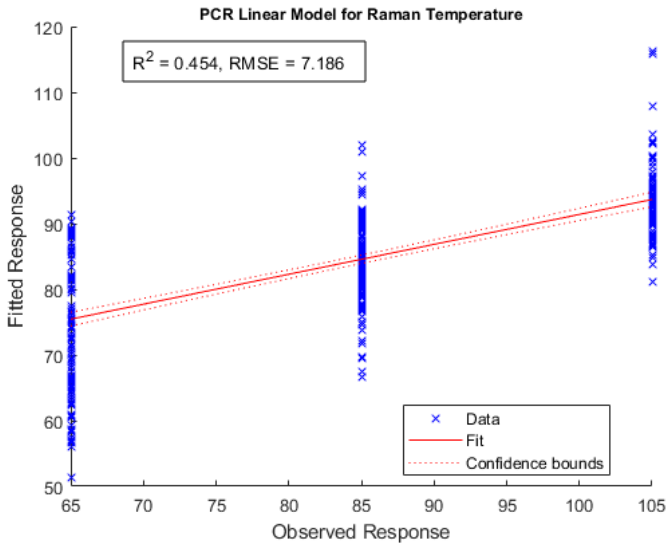

Figure 30: Raman Data PCR for Temperature

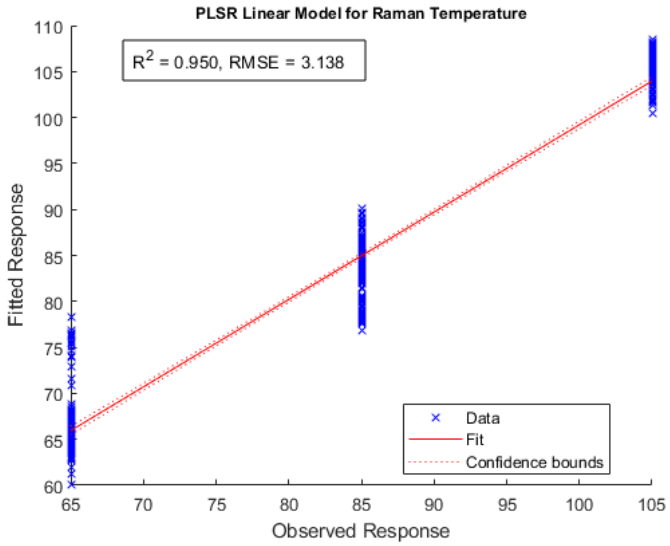

Figure 31: Raman Data PLSR for Temperature

Also, when comparing PCR to PLSR, both humidity and temperature have have better modeling when using PLSR over PCR by significant margins. PLSR has $R^2 = .983, .950$ versus PCR models with an $R^2 = .728, .454$ for humidity and temperature respectively. This due to the difference in the models techniques by which they reduce

the feature space and deal determining the separation between data points.

4.1.4 Raman Data Random Forest

In addition to PCR and PLSR, a Random Forest model was applied to the various response variables. The Random Forest model uses treatments categorical responses for classification, humidity, temperature, exposure time, and atmospheric type.

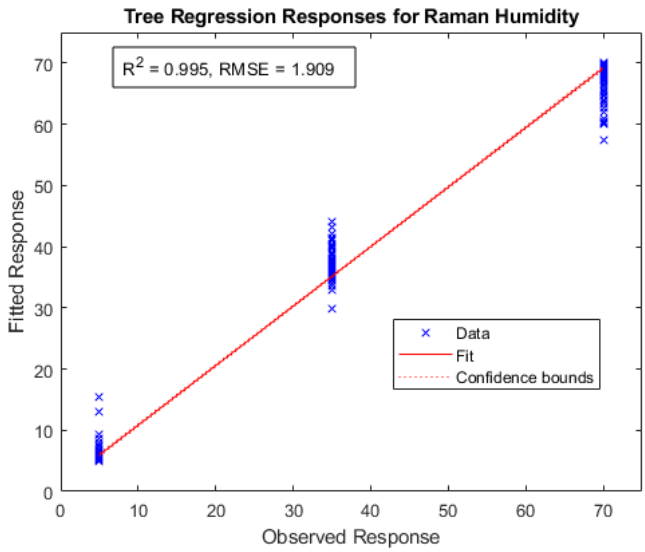

Figure 32: Raman Data Random Forest for Humidity Regression

The model predicted all categorical responses with 100% accuracy and recall, which can be seen in Appendix A. This is done using no training-test split for the data. Within the Random Forest model a 80/20% training-validation split when building the models. The separation seen in the PCA evaluation supports the performance of this modeling due to fact that there is generally good separation for each response variable.

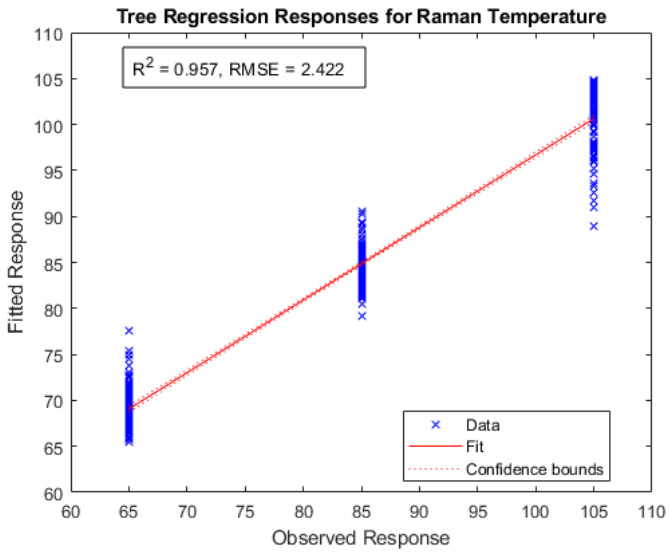

Figure 33: Raman Data Random Forest for Temperature Regression

In addition to the categorical models, regression models were built using Raman spectral data for humidity and temperature. In both cases Random Forest perform better than either PCR or PLSR, with an $R^2 = .995$ and $RMSE = 1.909$ for humidity and $R^2 = .957$ and $RMSE = 2.422$ for temperature, as shown in Figure 32 and Figure 33 respectively. These R^2 values are an average over 10 different Random Forest models. This was done due to the randomness by with the decision trees are built in the model, which creates a small variance in the R^2 and $RMSE$ values for each model built.

A training-test split is introduced to the Raman spectral data, to evaluate the overall performance that the models have in practical application to new data. Holding 20% of the data for a test set, 80% of the Raman spectral data is used to build categorical and regression models for the various response variables. In Figures 34 and 35 the performance for the model can be seen.

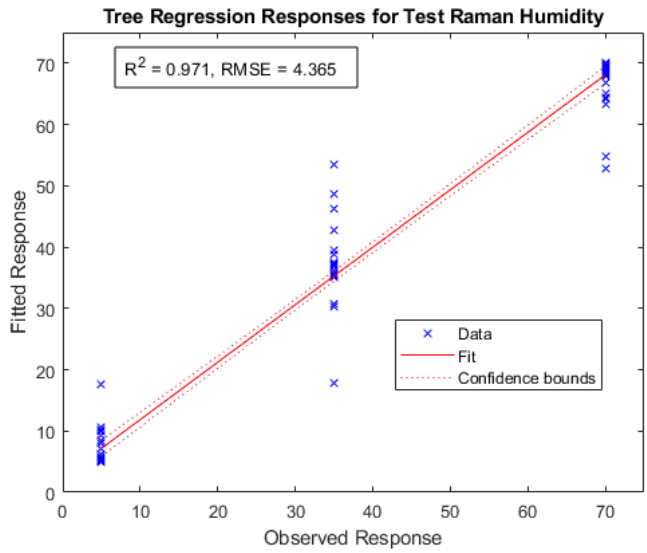

Figure 34: Raman Data Random Forest Model Test for Humidity Regression

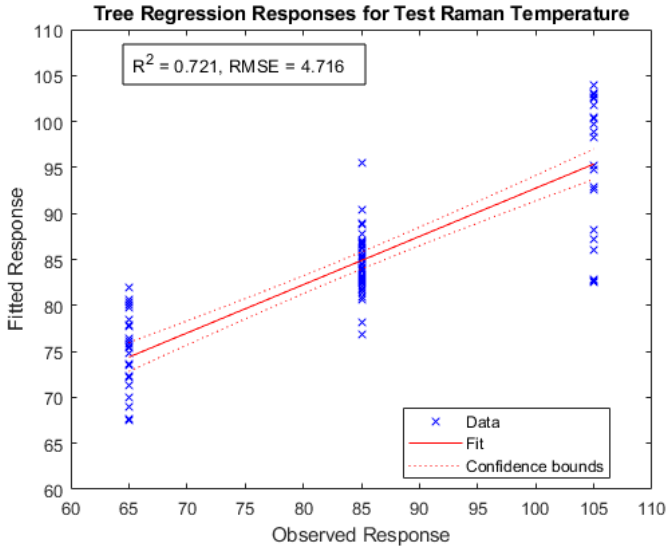

Figure 35: Raman Data Random Forest Model Test for Temperature Regression

For humidity the model performs well with an $R^2 = .971$ and $RMSE = 4.365$ and $R^2 = .721$ and $RMSE = 4.716$ for temperature. This is not unexpected with the analysis done into the component spaces for humidity and temperature where temperature was typically less separable.

Next, the categorical responses are tested to measure performance. This is done for classification, humidity, temperature, exposure time, and fill gas. for treatment classification the model performs well, 92% accuracy, but has a clear deficiency with some of the lower humidity and lower temperature treatments as seen in Figure 36.

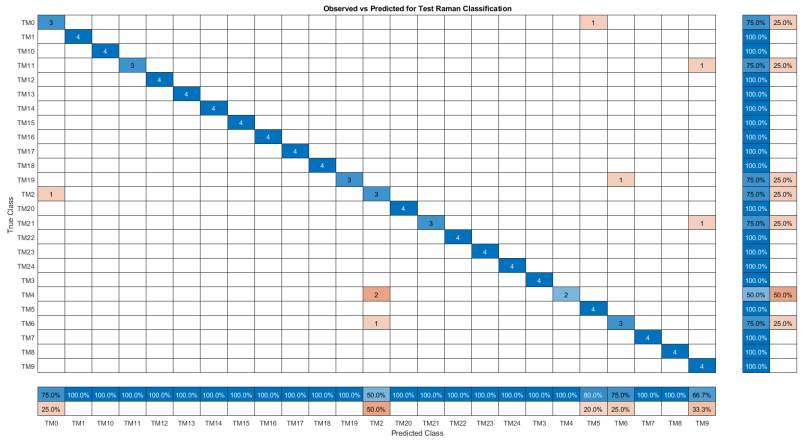

Figure 36: Raman Data Random Forest Model Test for Classification

For humidity, in Figure 37 did very well with only one error, or 99% accuracy, in the test data. However, for temperature the model performs slightly worse with a recall dropping to 70.8% and precision dropping to 80.4% as minimums, and a 85% accuracy, as shown in the confusion chart in Figure 38. When testing the fill gas response the accuracy is 91%. Considering the amount of nitrogen in air this is reasonable. This could be due the treatments with low reaction rates so that certain spectral peaks do not become prominent enough to have an effect.

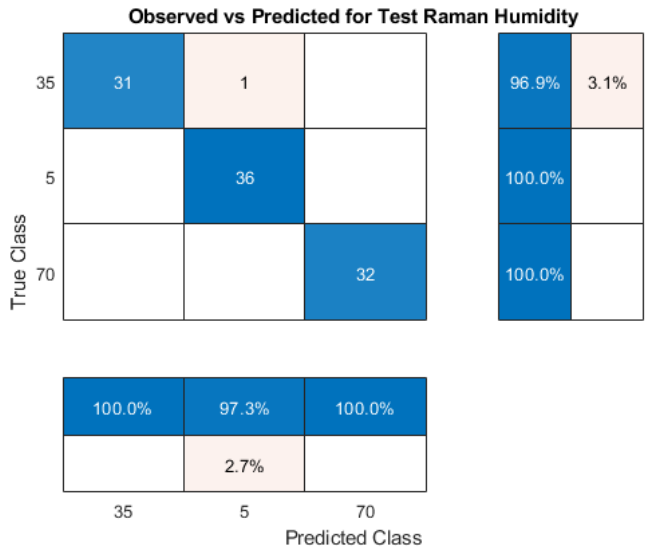

Figure 37: Raman Data Random Forest Model Test for Humidity

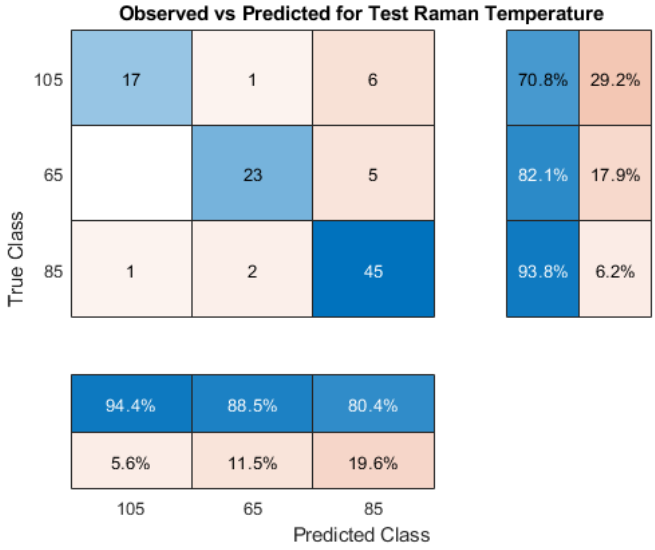

Figure 38: Raman Data Random Forest Model Test for Temperature

While having a 96% accuracy, exposure time modeling had issues specifically categorizing the 0hr treatments and the 2hr treatments, seen in Figure 40. This can be explained by the similarity seen in the PCA plots previously and the lack of separation in the short, lower humidity, lower temperature treatments and the 0hr

treatments. Overall, the Random Forest model does well with some distinctive areas of the data that cause issues when attempting to create models and predict responses.

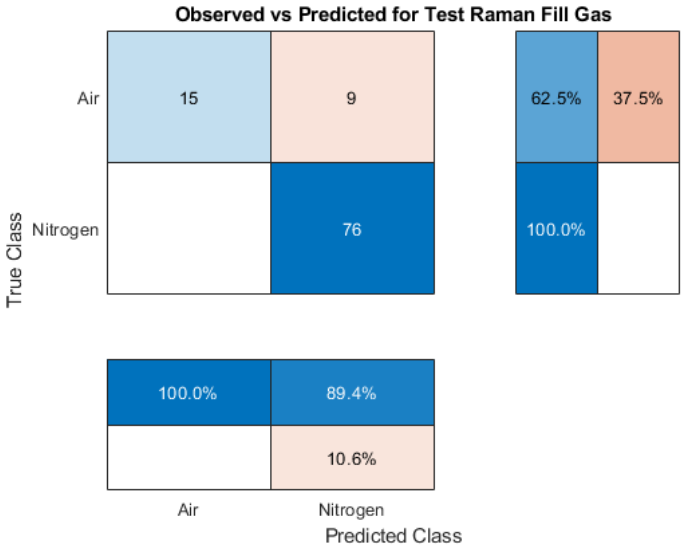

Figure 39: Raman Data Random Forest Model Test for Fill Gas

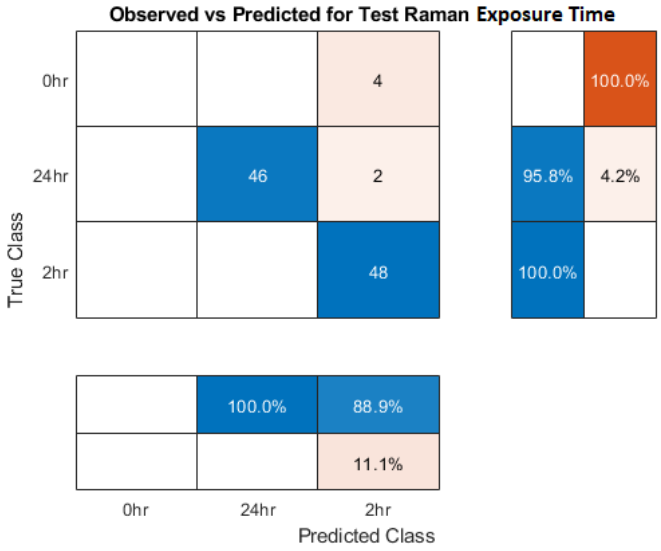

Figure 40: Raman Data Random Forest Model Test for Exposure Time

4.2 LIBS Data Analysis

Analysis is performed using only the normalized LIBS spectra using spectral analysis, PCA, PCR and PLSR, and the Random Forest modeling techniques.

4.2.1 LIBS Spectra

Using the main lines in Table 2.3.2, the LIBS spectra was investigated for visual trends during the environmental treatments. Higher temperature and humidity settings are viewed due to the more exaggerated changes between each spectra.

A comparison of Figures 43, 44, 45, show how the LIBS spectra can fluctuate in intensity greatly. Very little can be visibly deduced from the LIBS spectra due to the fact that atomic lines are typically are present but simply have varying intensities. One trend that was seen is a general decrease in the Li elemental lines as humidity and treatment time increased, as can be seen in the Figures 44 and 45. This could be due to the density of the sample or even the color of the being lighter is something that would affect this. Also due to the volume increase when LiH absorbs moisture the amount of Li that gets pressed in to the sample pellet will also decrease. These can be seen in Figures 41 and 42.

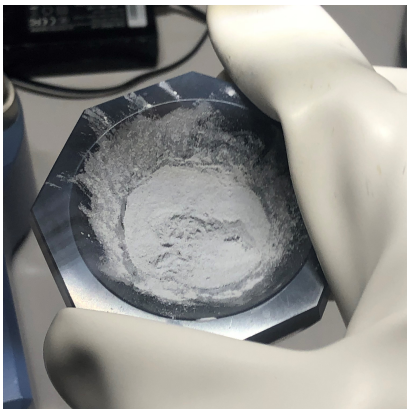

Figure 41: LiH Before Treatment

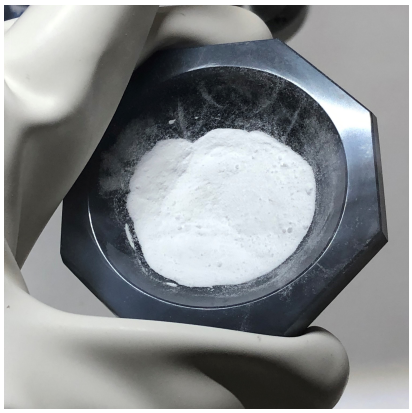

Figure 42: LiH After 24hr Treatment

Also, when the nitrogen purges were halted and air was the atmosphere, a clear decrease in the 6 nitrogen lines as seen in the spectra..

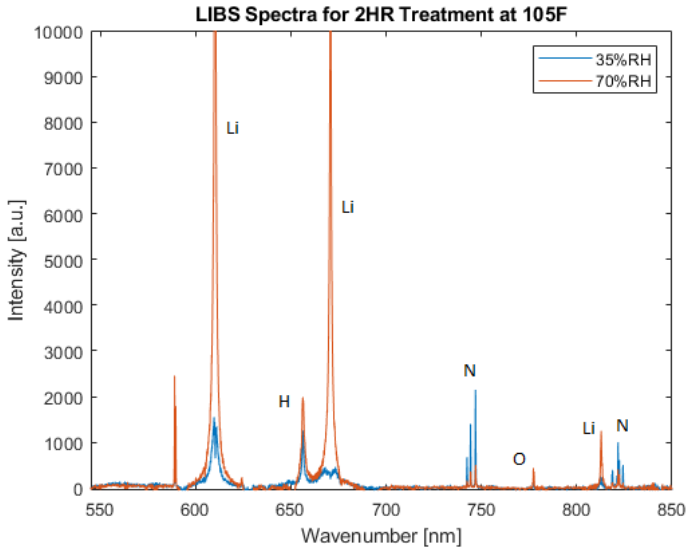

Figure 43: LIBS Spectra for 2hr Treatment

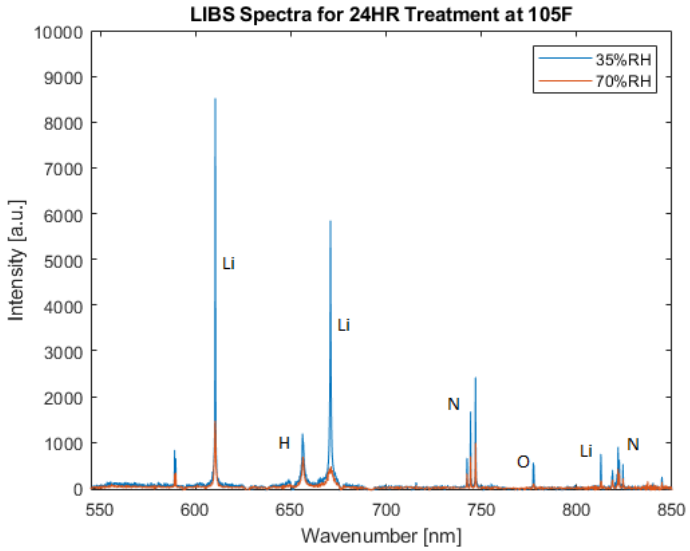

Figure 44: LIBS Spectra for 24hr Treatment

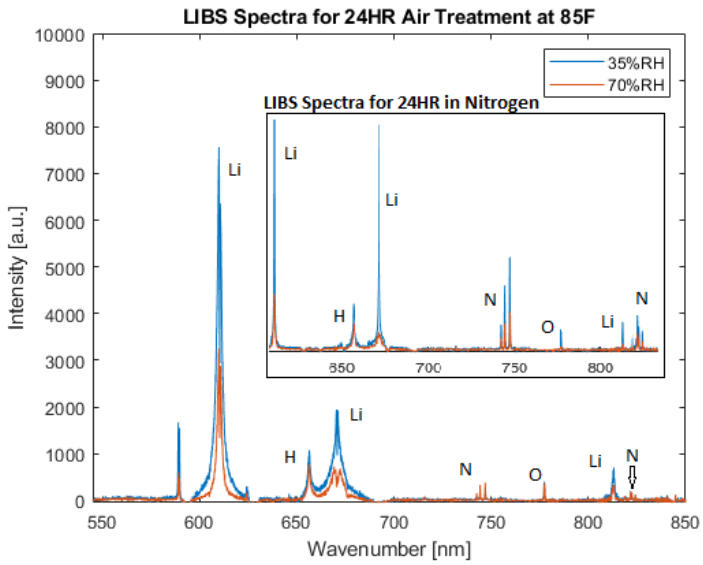

Figure 45: LIBS Spectra for 24hr Treatment in Air

It is much harder to make any judgement of trends occurring in the treatments, when compared to the Raman spectra due to the occurrence of most the significant emission lines in almost every spectra. The lack of certain elemental lines such as the Carbon line around 833.514 nm also exaggerates this issue. Only in the low humidity treatments when the oxygen lines and hydrogen lines are much lower in intensity is it possible to distinguish between the LiH and $LiOH$ for example.

4.2.2 LIBS Data PCA

Using the LIBS spectral data, PCA is used to provide a gauge of the separation in a component space and the viability of modeling using this data. This is done keeping the response variables of overall treatment classification, humidity, temperature, exposure time, and atmospheric type in mind. Treatment setting can be reference in Table 3.1.1 for clarity of these separations.

In general the spread of the LIBS data was greater in component space, and the class separation was less clear. In Figure 46, higher humidity treatments can be seen

more to the left (in green, red, and yellow), while lower humidity (in blue and pink) and some high humidity treatments with low exposure time (Orange and purple) are more to the right.

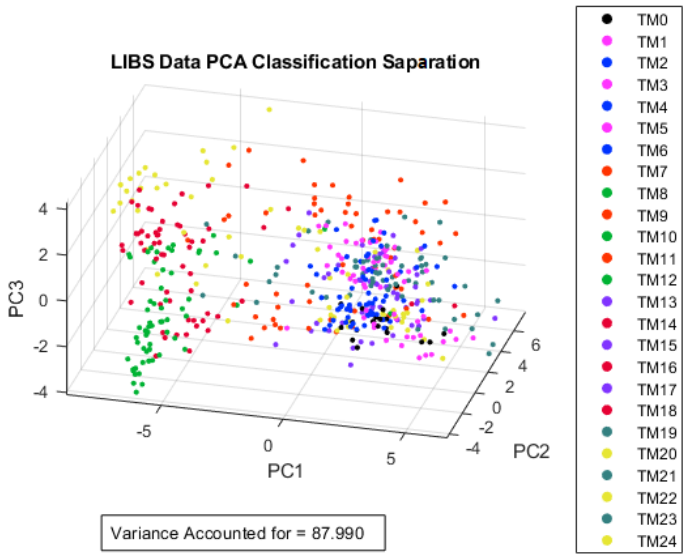

Figure 46: LIBS PCA for Classification Separation

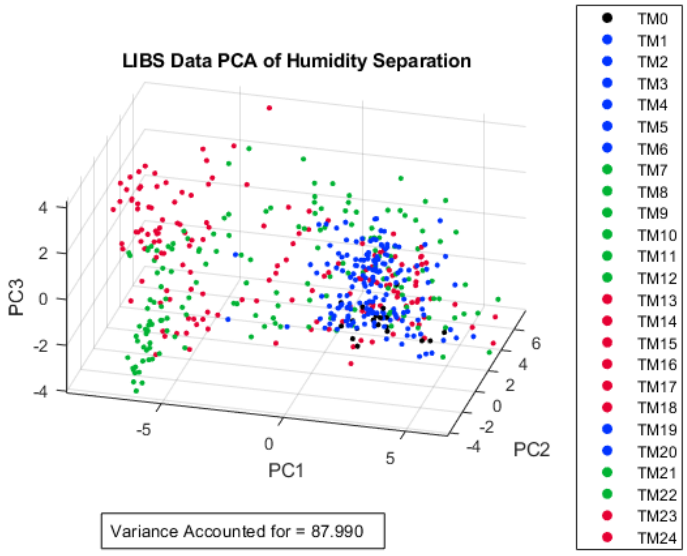

Figure 47: LIBS PCA for Humidity Separation

Observing Figures 47 and 48, the trend of higher humidity treatments being on the left (in green and red) and low humidity treatments are to the right is more visible. As well as the fact the in Figure 48 some of the longer exposures with lower humidity (in green) group with the shorter exposure treatments. Also, in each PCA figure the baseline treatments (in black), is grouped with low humidity and low exposure time treatments.

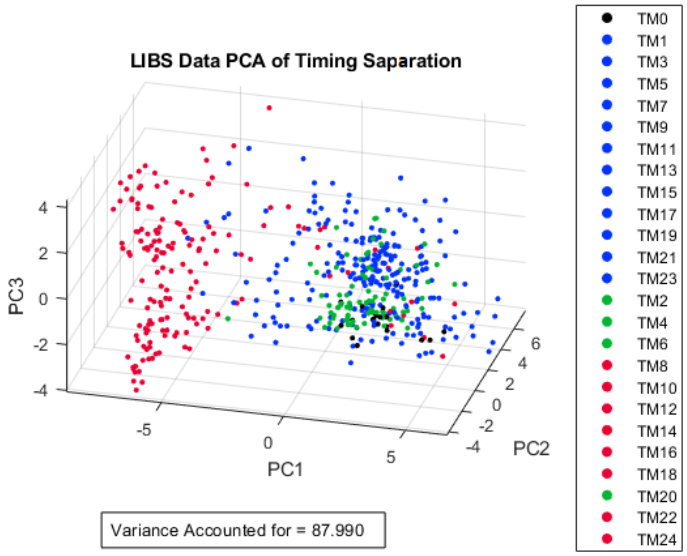

Figure 48: LIBS PCA for Exposure Timing Separation

4.2.3 LIBS Data PCR and PLSR

Using PCR and PLSR modeling, the LIBS spectral data was applied to humidity and temperature as the only two continuous responses focused on in this experiment.

4.2.3.1 LIBS Data PCR and PLSR for Humidity

How well PCR and PLSR performs with the LIBS data can be seen by viewing the first two components after feature reduction occurs in each model. In Figure 49 the overall lack of separation indicates that PCR for the LIBS data will perform worse

than PLSR when compared to Figure 50.

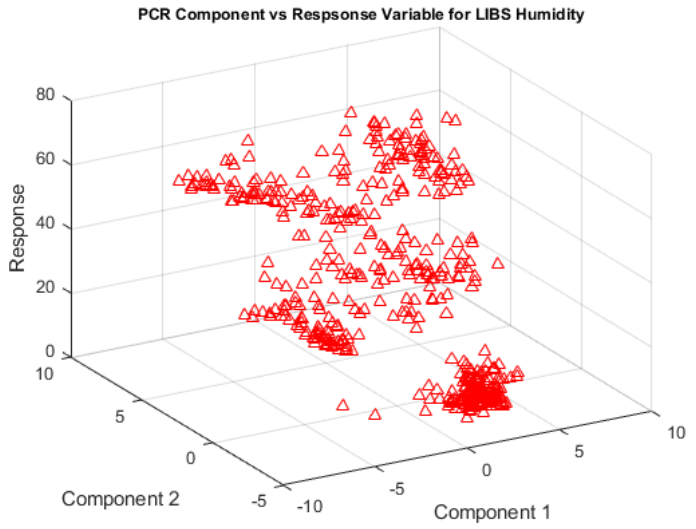

Figure 49: LIBS Data PCR Component Space Versus Humidity

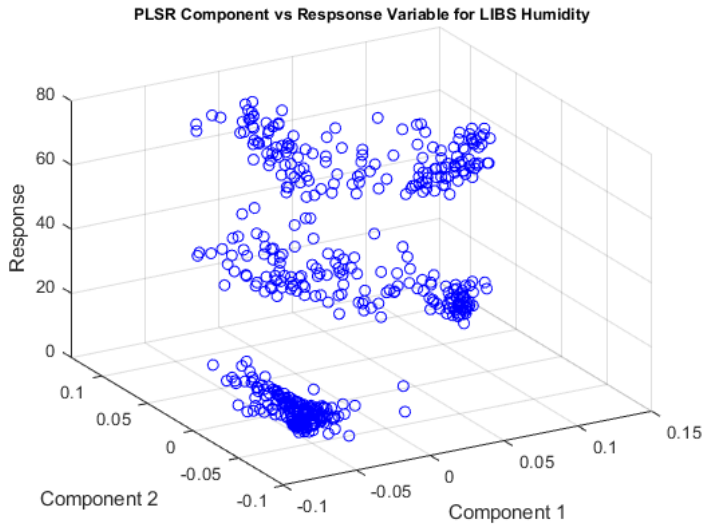

Figure 50: LIBS Data PLSR Component Space Versus Humidity

The PCR and PLSR models using LIBS data need approximately 15 components to account for 95% of the variance . This is less than the 25 needed for the Raman data, and thus 25 components are used to for the LIBS data modeling to obtain

comparable results. Also, the slower increase of variance accounted for per component indicates that PCR will be less accurate in comparison to PLSR.

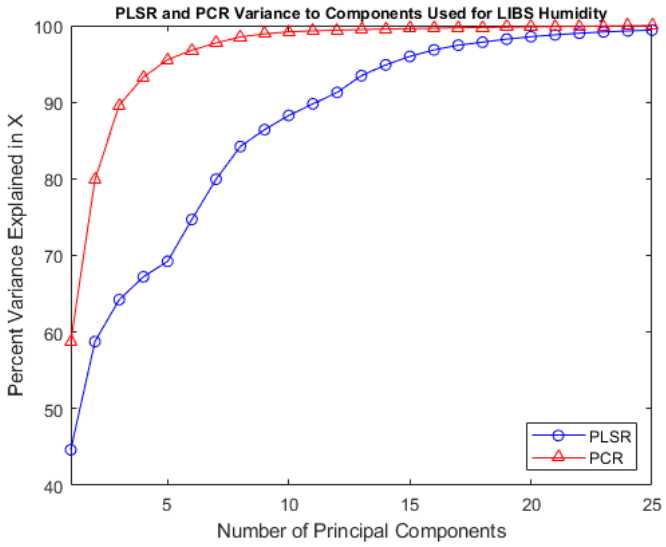

Figure 51: LIBS Data PCR and PLSR Variance Per Component for Humidity

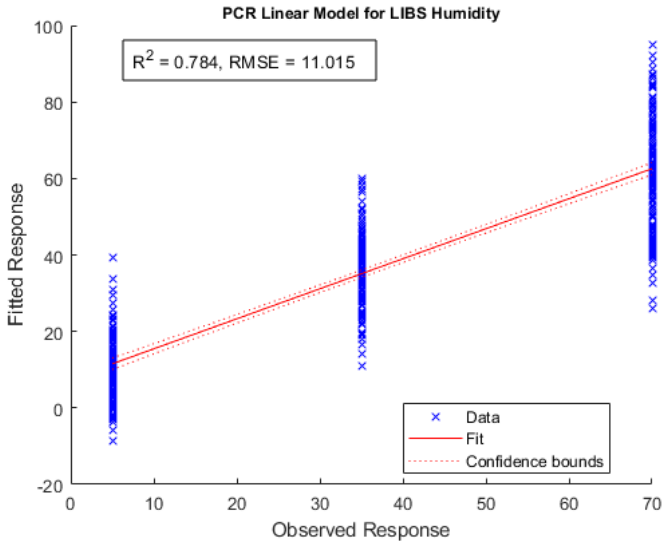

Figure 52: LIBS Data PCR for Humidity

As expected from the component space plots for PCR and PLSR, PCR performs worse with a $R^2 = .784$ and $RMSE = 11.015$ seen in Figure 52. PLSR performs well

with a $R^2 = .994$ and $RMSE = 2.055$, seen in Figure 53. Overall the LIBS data modeling does quite well the the humidity response variable.

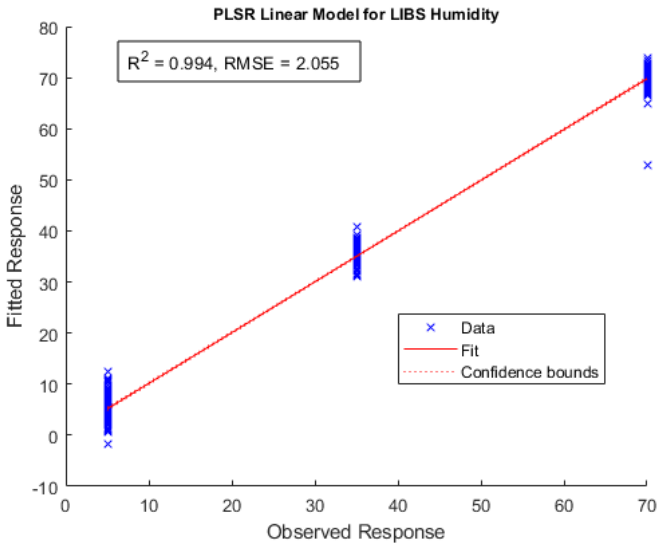

Figure 53: LIBS Data PLSR for Humidity

4.2.3.2 LIBS Data PCR and PLSR for Temperature

PCR and PLSR was then applied to the temperature responses. When viewing the component space for PCR and PLSR with the temperature response variable, the separation in between the data is not distinct. While the PLSR component space in Figure 55 still has some slight separation, the PCR component space almost has poor separation as seen in Figure 54.

By looking at the variance accounted for in Figure 56, a number of components needed to reach 95% variance was used. Also, the similar to the humidity case, the amount of variance per component is smaller for PLSR than PCR pointing to a better modeling ability for PLSR.

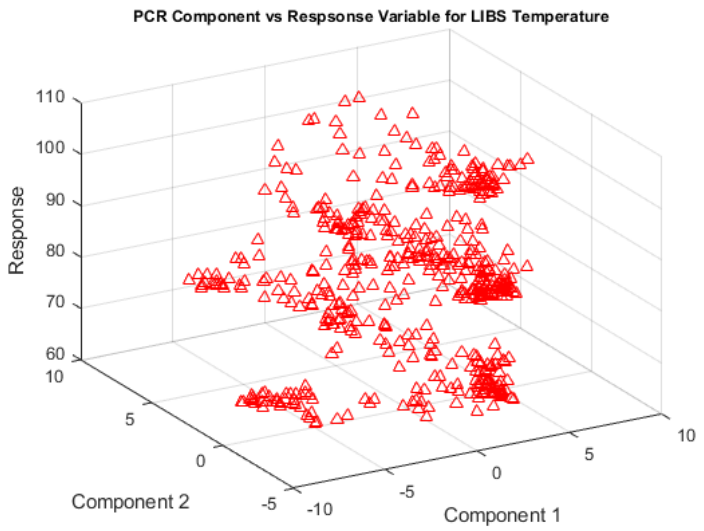

Figure 54: LIBS Data PCR Component Space Versus Temperature

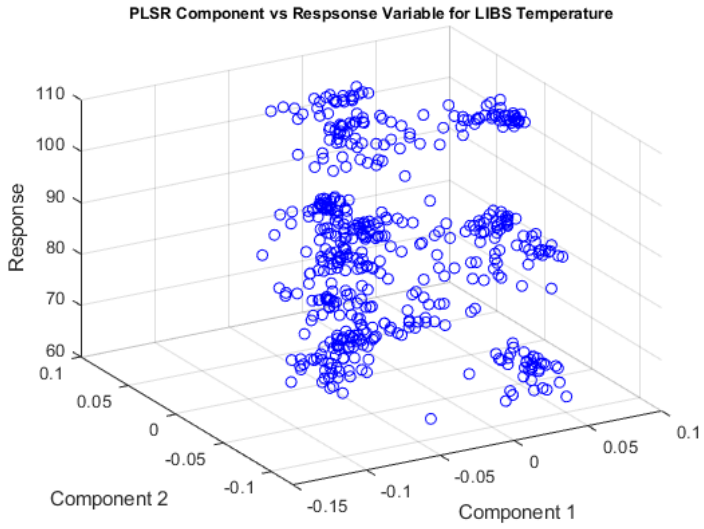

Figure 55: LIBS Data PLSR Component Space Versus Temperature

As indicated by the component space plots and variance plots, PLSR performs better than PCR with a $R^2 = .977$ and $RMSE = 2.144$, seen in Figure 58. PCR performs worse with a $R^2 = .370$ and $RMSE = 6.971$ as seen in Figure 57. Unlike with the PCR for humidity using LIBS data, PCR for temperature performs poorly.

PLSR for temperature does perform worse than for humidity, but still works very well for the LIBS data.

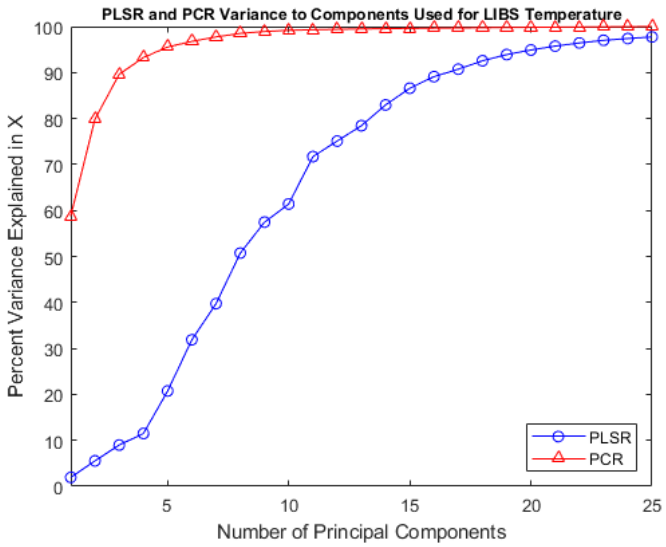

Figure 56: LIBS Data PCR and PLSR Variance Per Component for Temperature

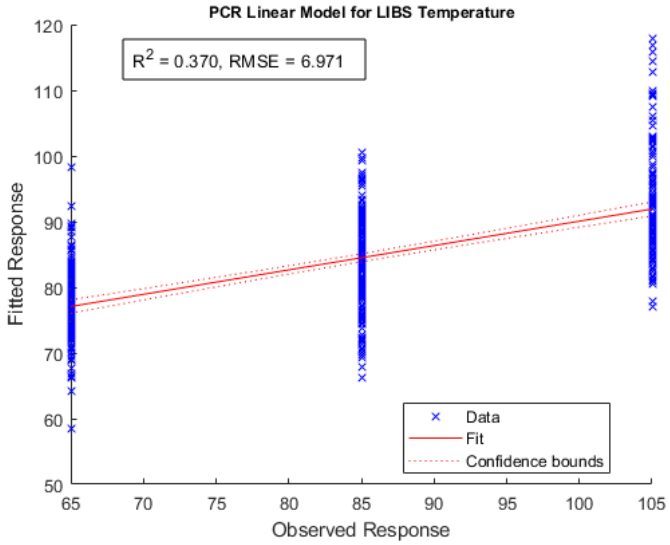

Figure 57: LIBS Data PCR for Temperature

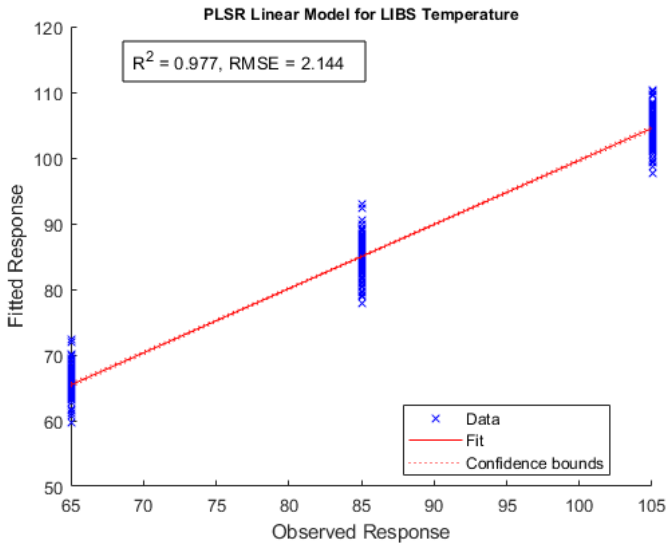

Figure 58: LIBS Data PLSR for Temperature

4.2.4 LIBS Data Random Forest

The Random Forest model used for treatments of classification, predict 100% accuracy and recall. Also, Random Forest models predicted 100% accuracy and recall for humidity, temperature, exposure time, and atmospheric type. These confusion charts for the models performance when using categorical response variables can be viewed In Appendix A. Due to fact that there is generally good separation for each response variable, as seen in the PCA plots, the Random Forest model works very well on the categorical responses.

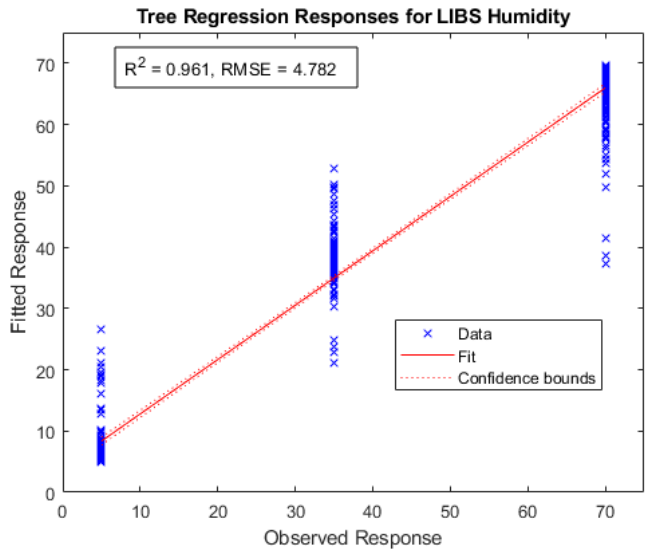

Figure 59: LIBS Data Random Forest for Humidity Regression

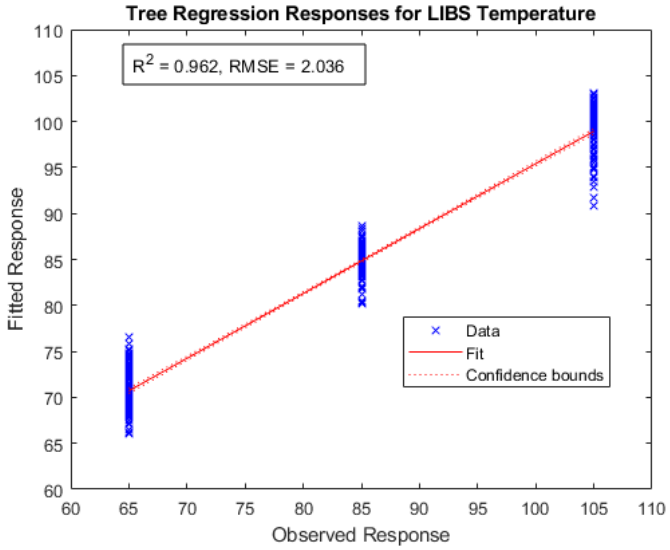

Figure 60: LIBS Data Random Forest for Temperature Regression

Random Forest regression models were built for both humidity and temperature using the LIBS spectral data. Random Forest performs only slighty better for temperature than for humidity with a $R^2 = .962$ and $RMSE = 2.036$, and $R^2 = .961$ and $RMSE = 4.782$. This is seen in Figures 59 and 60. for the LIBS spectral data, PLSR

performs better than PCR or the Random Forest model when looking at temperature.

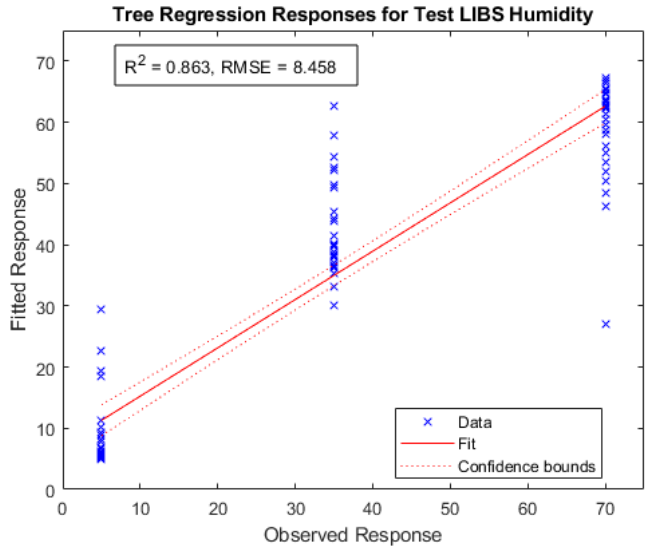

Figure 61: LIBS Data Random Forest Model Test for Humidity Regression

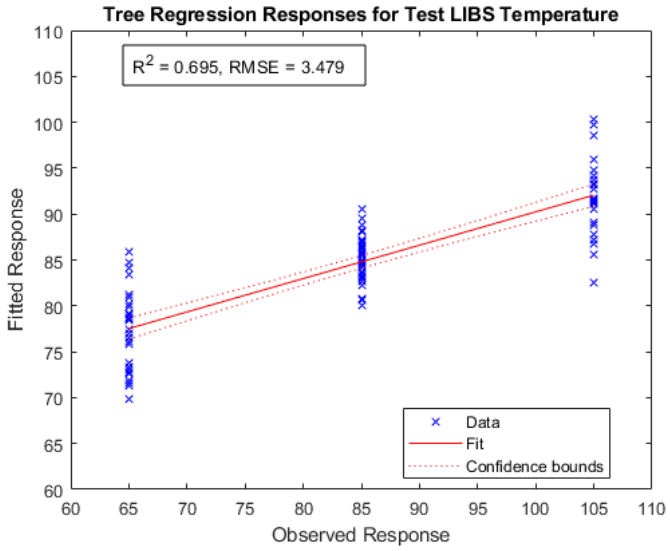

Figure 62: LIBS Data Random Forest Model Test for Temperature Regression

Just as for the Raman data, a training-test split is introduced to the LIBS spectral data. This is done to evaluate the overall performance that the models have in practical application to new data. Holding 20% of the data for a test set, 80% of the

LIBS spectral data is used to build categorical and regression models for the various response variables. In Figures 61 and 62 the performance for the model can be seen.

For humidity the model performs reasonably well with a $R^2 = .863$ and $RMSE = 8.458$ and $R^2 = .695$ and $RMSE = 3.479$ for temperature. This is not unexpected with the analysis done into the component spaces for humidity and temperature where temperature was typically less separable.

Next, the categorical responses are tested to measure performance. This is done for classification, humidity, temperature, exposure time, and fill gas. for treatment classification the model performs decently well, with 90% accuracy, but has a issue distinguishing treatments with similar conditions in some cases. Most of the incorrect classifications seem as though they are due to that lack of clear separation between similar treatments as shown in Figure 63.

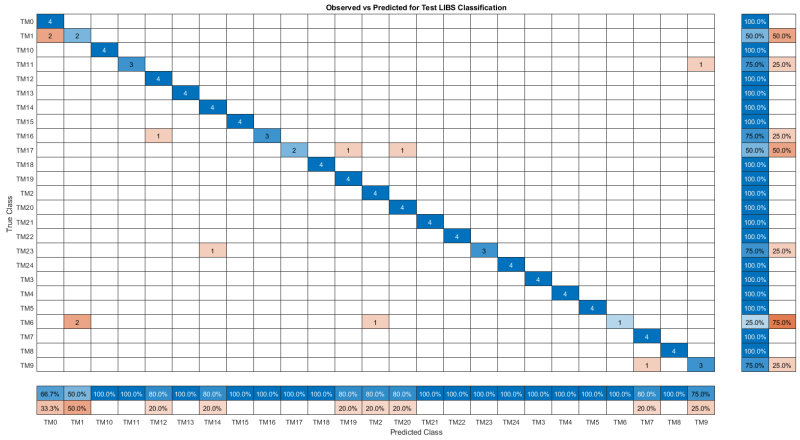

Figure 63: LIBS Data Random Forest Model Test for Classification

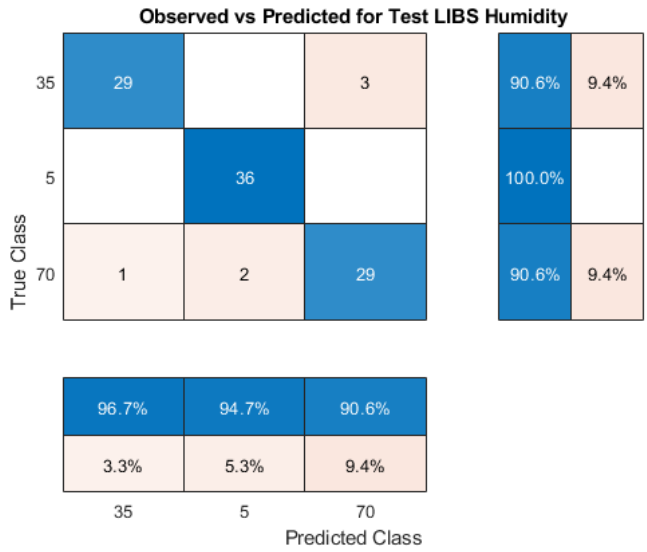

Figure 64: LIBS Data Random Forest Model Test for Humidity

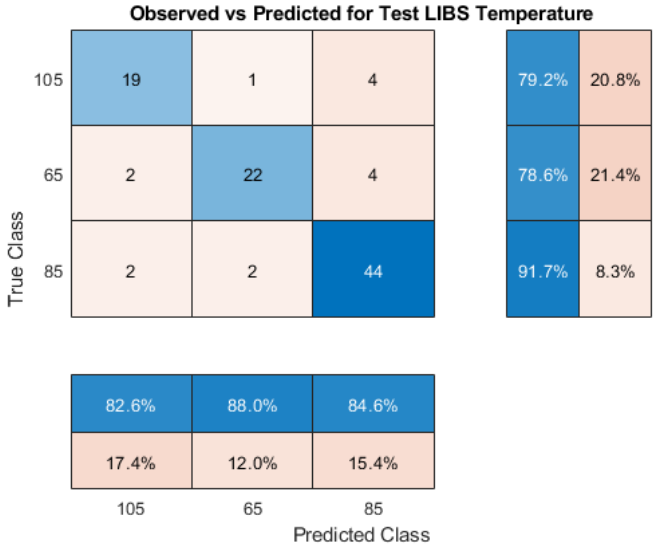

Figure 65: LIBS Data Random Forest Model Test for Temperature

In Figure 64, humidity modeling does very well with a 94% accuracy in the test model responses. However, for temperature the model performs slightly worse with a recall dropping to 78.6% and precision dropping to 82.6% as minimums, and a 85% accuracy. This is shown in the confusion chart in Figure 65. When testing the fill

gas response, the model does fairly well with an accuracy of 95%. Which considering the amount of nitrogen in air and the ability of LIBS to see the atomic nitrogen lines makes sense.

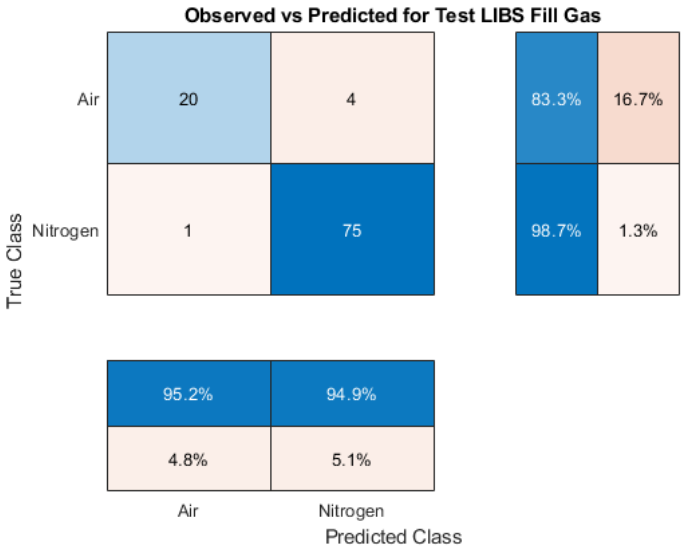

Figure 66: LIBS Data Random Forest Model Test for Fill Gas

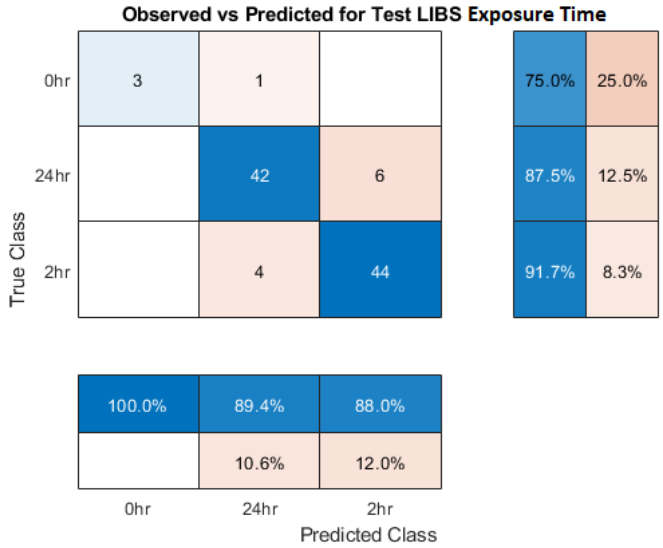

Figure 67: LIBS Data Random Forest Model Test for Exposure Time

As shown in Figure 67, the model does reasonably well with exposure time pre-

diction. The recall of 75% for the 0hr treatments is likely due to lack of data point for this type of treatment. This model performs decently overall with an accuracy of 86%. Generally the Random Forest model does very well when using LIBS data, save for the temperature and exposure time response variables.

4.3 Raman and LIBS Data Fusion Analysis

Using the main lines in Table 2.3.2 and Table 2.3.1, the normalized Raman and LIBS spectra are appended to eachother. This Raman and LIBS fusion data is then analysed with, PCA, PCR and PLSR, and the Random Forest techniques.

4.3.1 Data Fusion PCA

In the component space of PCA, visible trends with classification separation, humidity, and exposure time are seen. Figure 68 shows clear separation between some treatments (in red and blue) with high and low humidity.

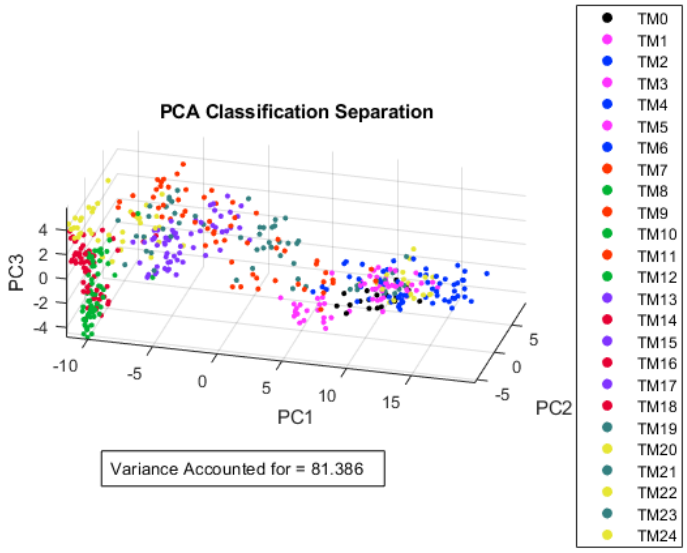

Figure 68: Data Fusion for Classification Separation

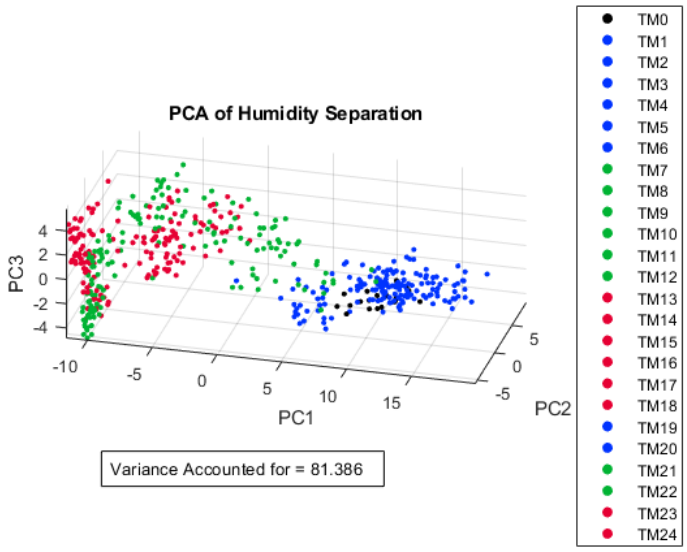

Figure 69: Data Fusion for Humidity Separation

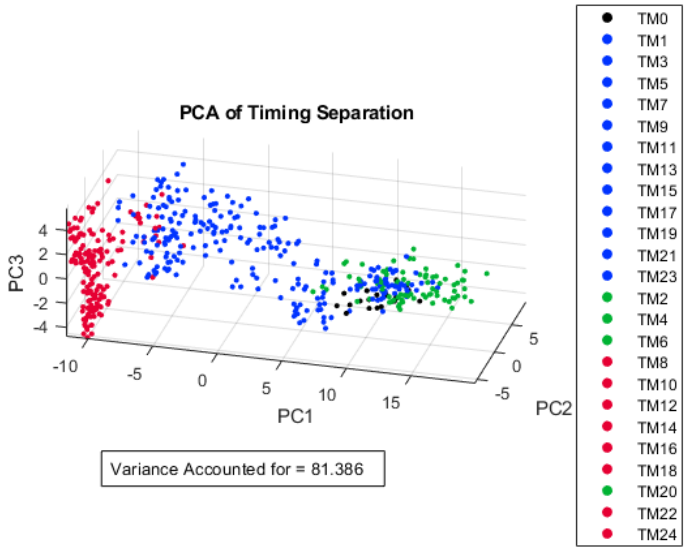

Figure 70: Data Fusion for Exposure Timing Separation

Some treatments have more overlap, such as treatments with low exposure time (in purple and orange). Using Figure 69 similar trends can be seen. Medium and high

humidity's (in green and red) have a great separation from the data for low humidity treatments (in blue).

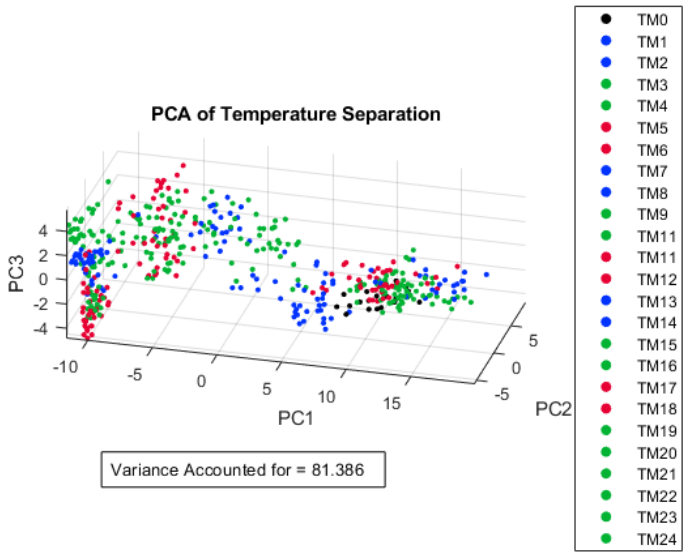

Figure 71: Data Fusion for Temperature Separation

There is also a clear mix of Medium and high humidity data as well. When looking a exposure time it becomes clear that the mixed data is mainly data of higher humidity but with low exposure time (in blue) as seen in Figure 70. Also it is seen that some of the longer treatments group with lower humidity treatments as well. When reaction rate of the LiH are considered this is understandable, due to the extreme dependence on humidity. Another notable feature can be seen in Figure 71, when is become clear that temperature is the most convoluted of the response variables in this study.

4.3.2 Data Fusion PCR and PLSR

Using PCR and PLSR modeling, the data fusion spectral data was applied to humidity and temperature as the only two continuous responses focused on in this

experiment.

4.3.2.1 Data Fusion PCR and PLSR for Humidity

As with the other models, how PCR and PLSR will perform with the data fusion can be seen by viewing the first two components after feature reduction occurs in each model. In Figures 72 and 73, there seem to be great separation in the component space for PLSR and also good separation for PCR. Both models are expected to work well in the humidity case.

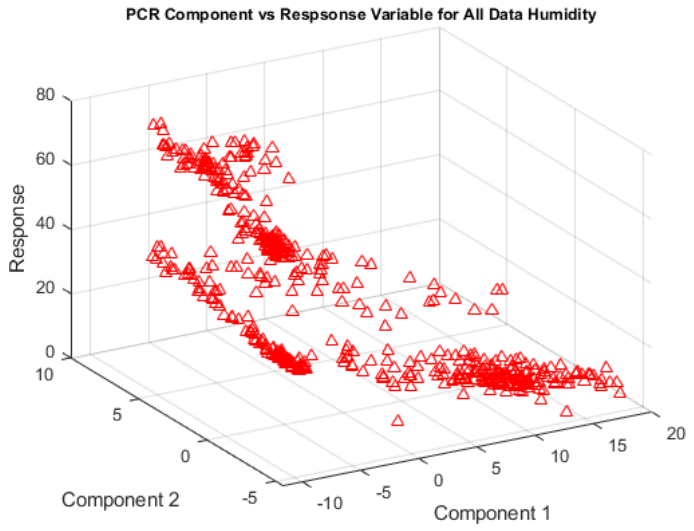

Figure 72: Data Fusion PCR Component Space Versus Humidity

When looking at Figure 74, only about four components are need for PCR to reach 95% variance, while PLSR needs closer to 16. However, 25 components are used to accommodate comparisons to previous modeling. The slower increase of variance per component indicates that PLSR will work better than PCR.

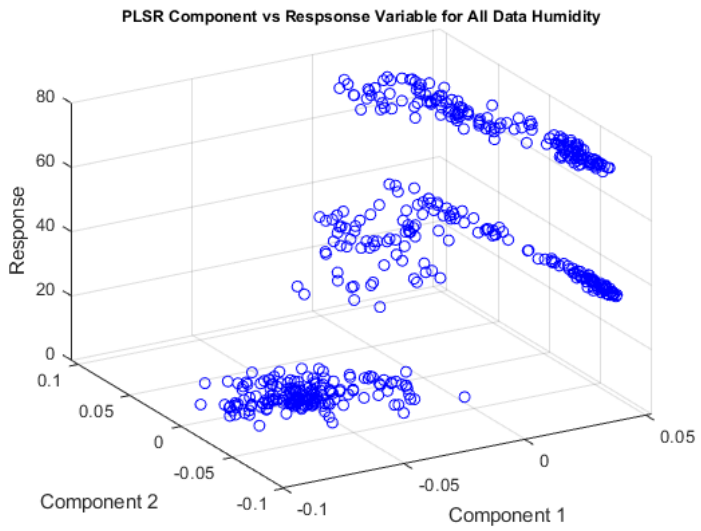

Figure 73: Data Fusion PLSR Component Space Versus Humidity

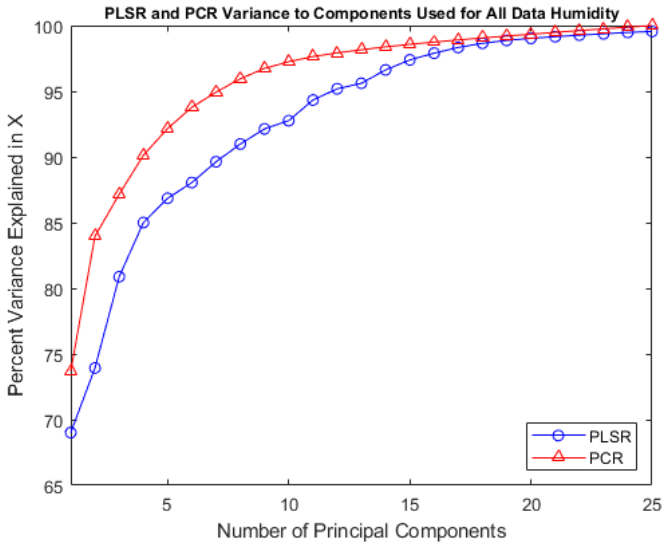

Figure 74: Data Fusion PCR and PLSR Variance Per Component for Humidity

PLSR does perform better than PCR, with an $R^2 = .995$ and $RMSE = 1.796$ and $R^2 = .860$ and $RMSE = 9.303$ for humidity, as shown in Figure 76 and Figure 75 respectively. Using the data fusion works very well with PCR and PLSR for humidity modeling.

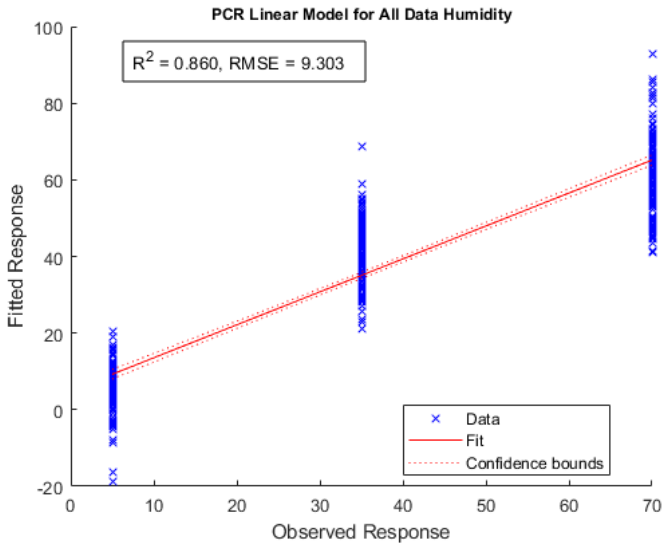

Figure 75: Data Fusion PCR for Humidity

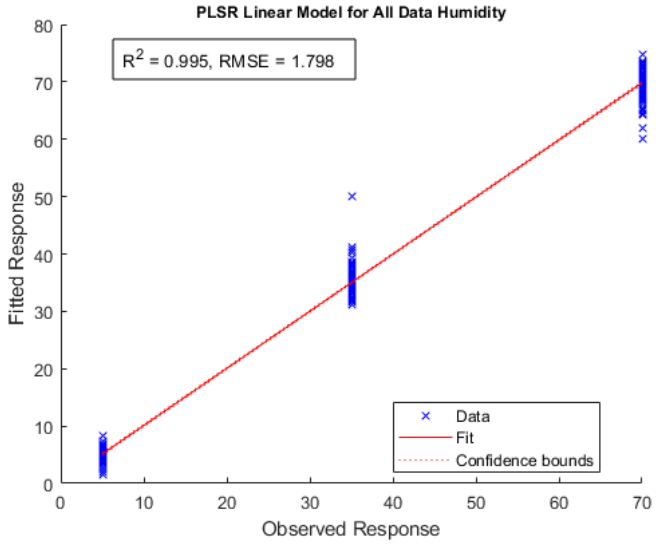

Figure 76: Data Fusion PLSR for Humidity

4.3.2.2 Data Fusion PCR and PLSR for Temperature

Just as will humidity, PCR and PLSR component space is evaluated to provide an indication of how well the data fusion performs after feature reduction with the

temperature response variable. In Figures 77 and 78, there is some separation in the component space for PLSR and less separation for PCR.

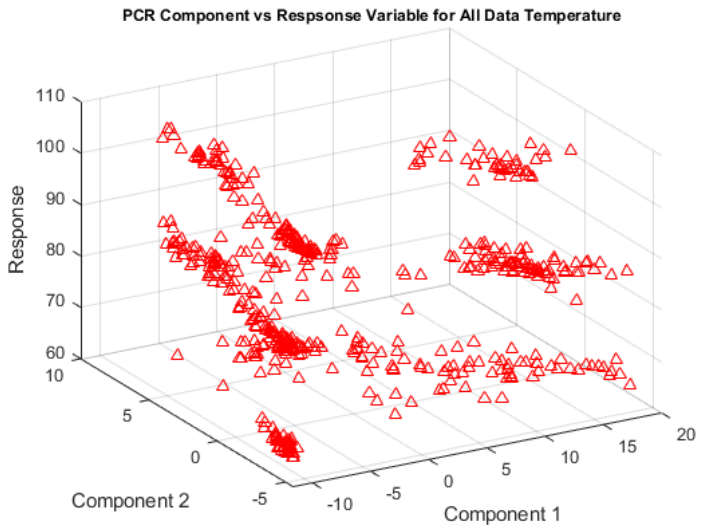

Figure 77: Data Fusion PCR Component Space Versus Temperature

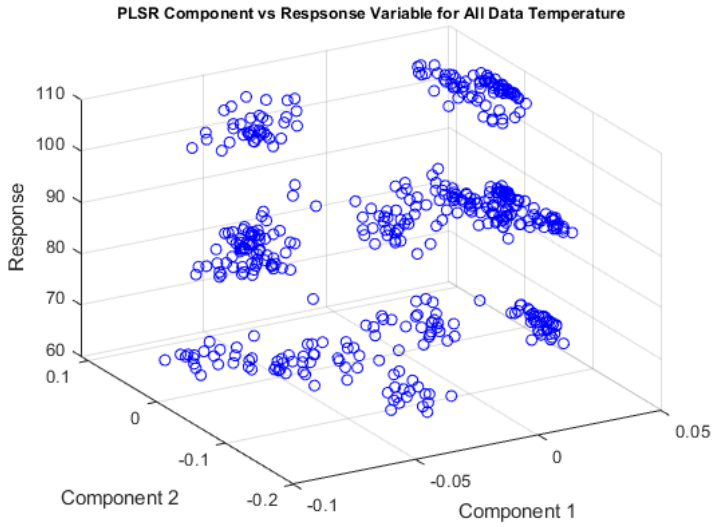

Figure 78: Data Fusion PLSR Component Space Versus Temperature

The variance accounted for in the PCR reaches 95% at approximately 4 components used as seen in Figure 79. At the same time, PLSR needs approximately 18

components to reach the same variance. Again, 25 components are used as in other modeling. PCR uses very few components to reach high variance while PLSR needs more, pointing to PLSR being able to more accurately model the data.

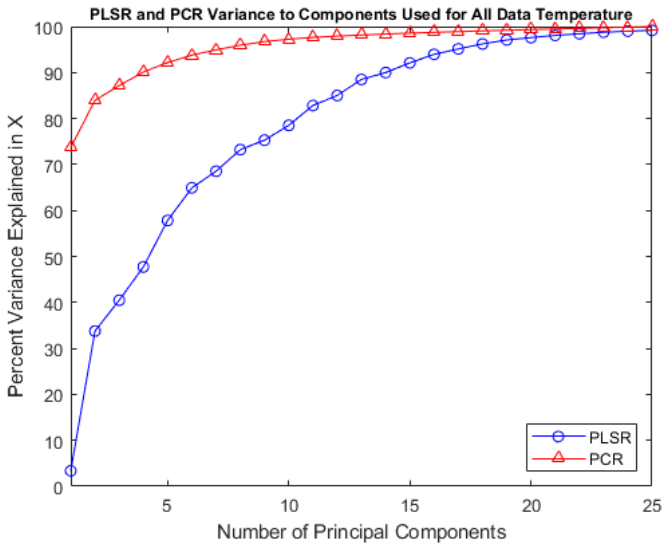

Figure 79: Data Fusion PCR and PLSR Variance Per Component for Temperature

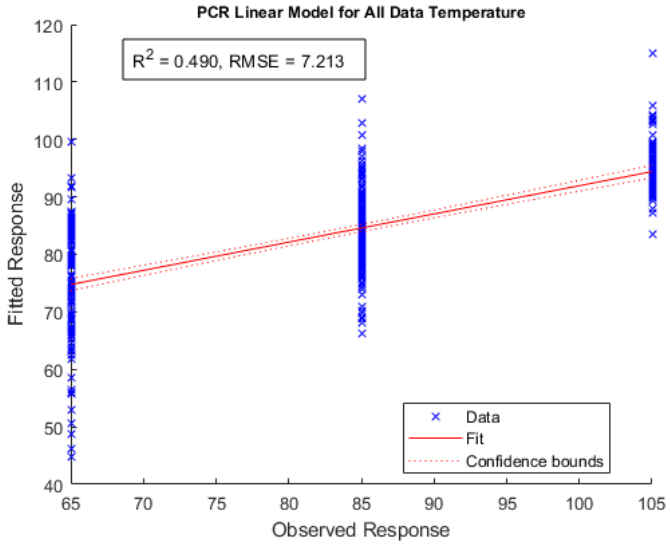

Figure 80: Data Fusion PCR Temperature

As expected PCR performs worse than PLSR, with an $R^2 = .490$ and $RMSE =$

7.213 and $R^2 = .992$ and $RMSE = 1.307$ for temperature, as shown in Figure 80 and Figure 81 respectively. Thus, for temperature responses PLSR works very well, while PCR works very poorly.

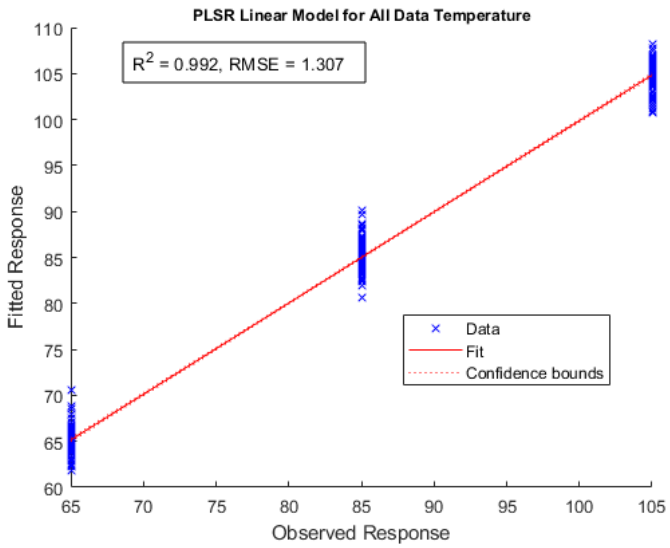

Figure 81: Data Fusion PLSR for Temperature

4.3.3 Fusion Data PCR and PLSR for Test Set

A training-test split was performed using an 80/20% split, for Fusion spectral data. Figures 82 and 83 show this data. For the training-test split, PLSR still out performs PCR when humidity is considered.

For humidity PLSR has an $R^2 = 1.000$ versus an $R^2 = .864$ for PCR. For the temperature response PLSR out performs PCR with $R^2 = .999$ versus $R^2 = .527$ respectively. These results are seen in Figures 84 and 85.

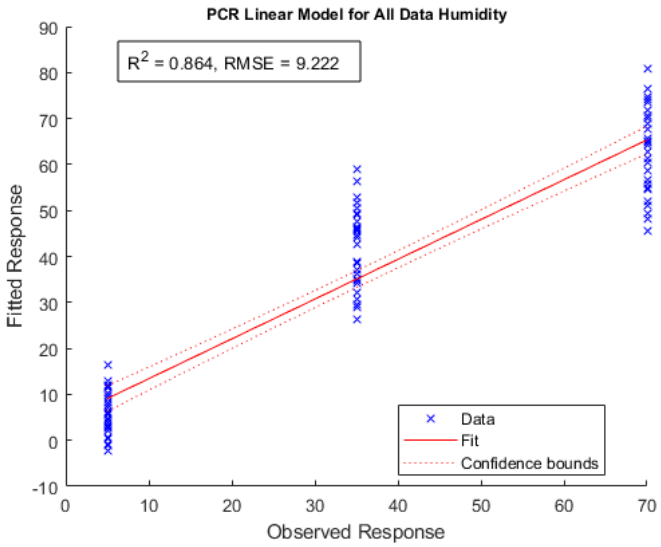

Figure 82: Fusion Data PCR Test for Humidity

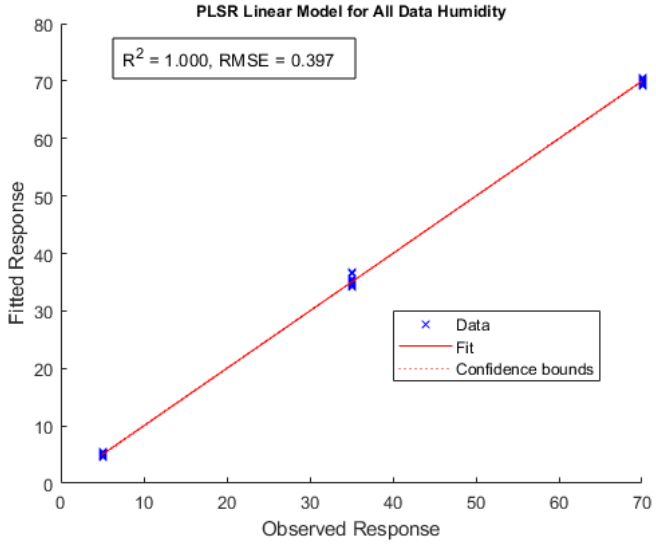

Figure 83: Fusion Data PLSR Test for Humidity

Keeping consistent with the trends seen in previous modeling, PLSR has better performance than PCR, in both humidity and temperature cases. When comparing humidity and temperature modeling performance, PLSR models have an $R^2 = 1.000$ and $R^2 = .999$. Thus, temperature continues to to the more difficult model to model.

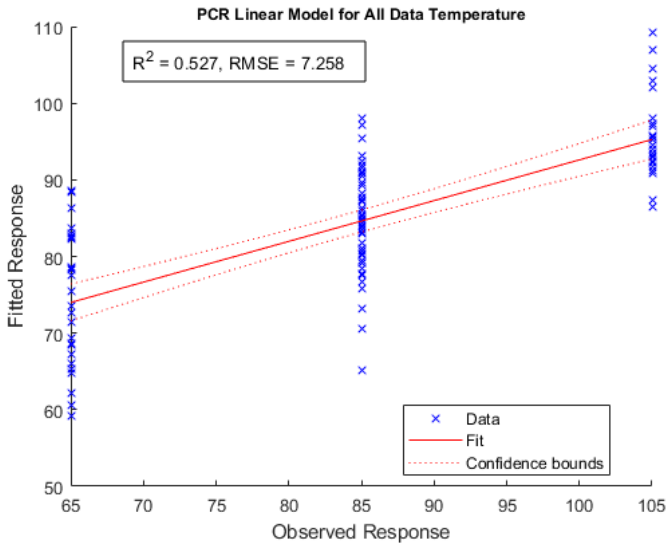

Figure 84: Fusion Data PCR Test for Temperature

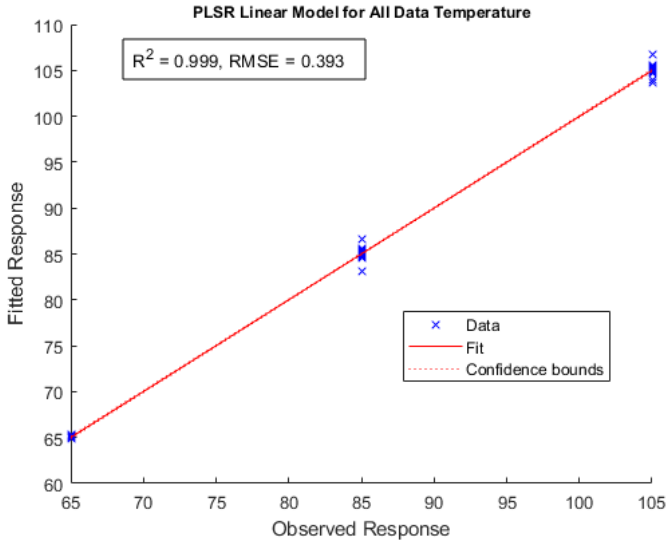

Figure 85: Fusion Data PLSR Test for Temperature

4.3.4 Data Fusion Random Forest

The Random Forest model is applied using treatments categorical responses for classification, humidity, temperature, exposure time, and atmospheric type as well

as continuous responses for humidity and temperature. The model predicted all categorical responses with 100% accuracy and recall, which can be seen in C. This is done using no training-test split for the fused data. for the all responses, Random Forest also uses a 80/20% training-validation split for additional model testing.

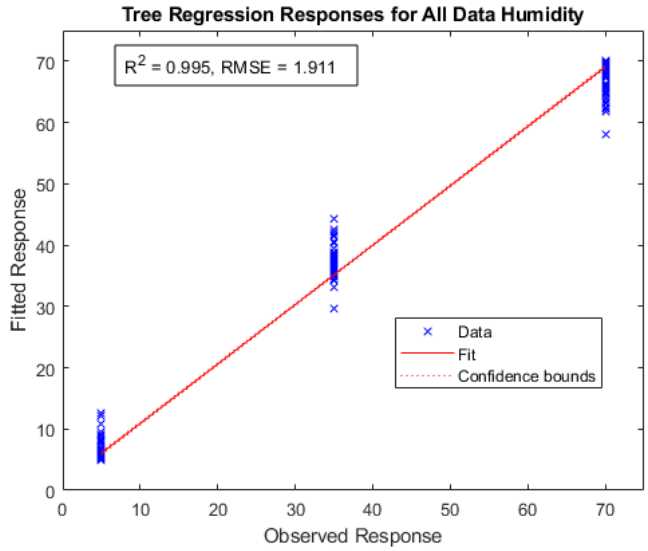

Figure 86: Data Fusion Random Forest for Humidity Regression

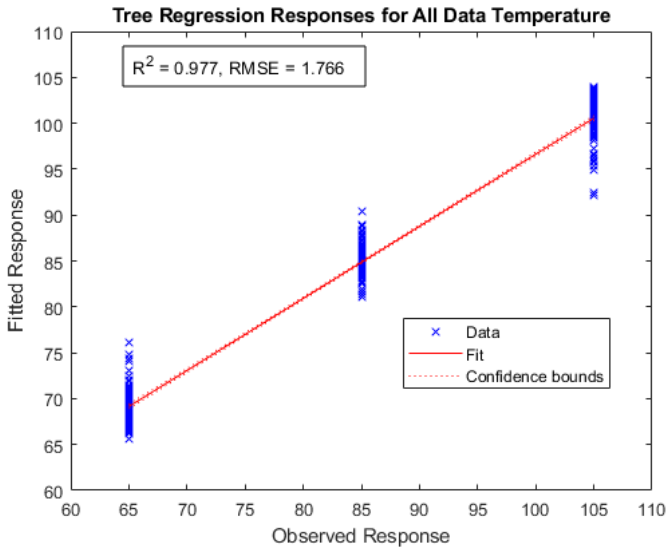

Figure 87: Data Fusion Random Forest for Temperature Regression

For humidity and temperature the Random Forest model is performed, and in both cases if performs very well. As shown in Figures 86 and 87, Random Forest modeling has a $R^2 = .979$ and $RMSE = 3.694$ for humidity and a $R^2 = .962$ and $RMSE = 2.053$ for temperature. These R^2 and $RMSE$ values are an average of 10 different models built and studied. Overall the fusion data is handled very well by the Random Forest model.

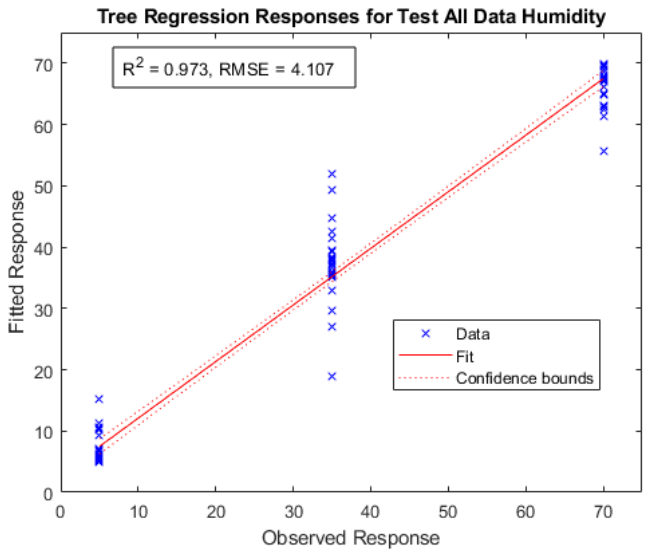

Figure 88: Data Fusion Random Forest Model Test for Humidity Regression

For the test data withheld from the fusion data, models are built using Random Forest and then tested with the test fusion data. These models have a $R^2 = .973$ and $RMSE = 4.107$ for humidity and a $R^2 = .813$ and $RMSE = 3.258$ for temperature. As shown in Figures 86 and 86, these models handle the individual responses variables quite well when presented with new data.

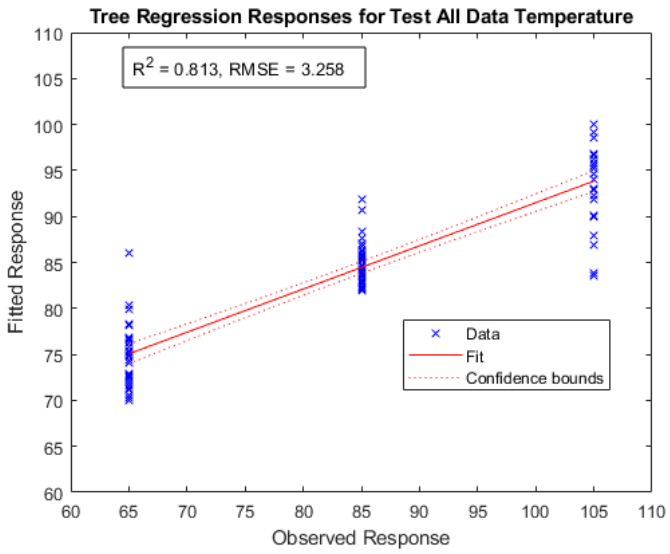

Figure 89: Data Fusion Random Forest Model Test for Temperature Regression

The categorical responses are tested to measure performance with classification, humidity, temperature, exposure time, and fill gas responses. for treatment classification the model performs very well, 97% accuracy, with only a few errors of similar treatments being classified to each other as seen in Figure 90.

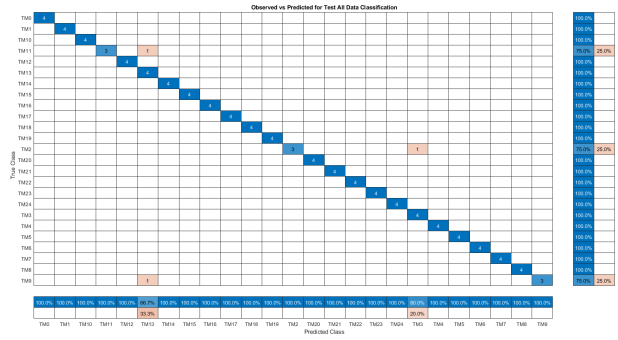

Figure 90: Data Fusion Random Forest Model Test for Classification

For the case of humidity, the Random Forest model accurately predicts each data set correctly to its corresponding response variable with no errors or 100% accuracy. However, for temperature the model does have a few errors leading to a recall min-

imum of 89.3% and a minimum precision of 88% and accuracy of 92%. These cases are shown in Figures 91 and 92

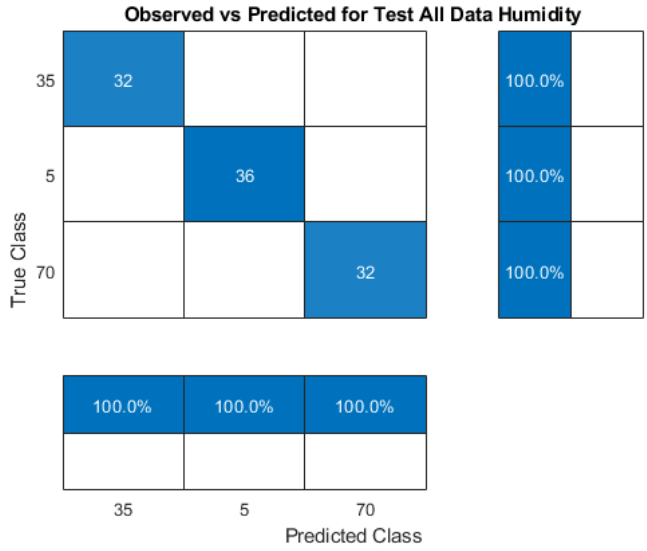

Figure 91: Data Fusion Random Forest Model Test for Humidity

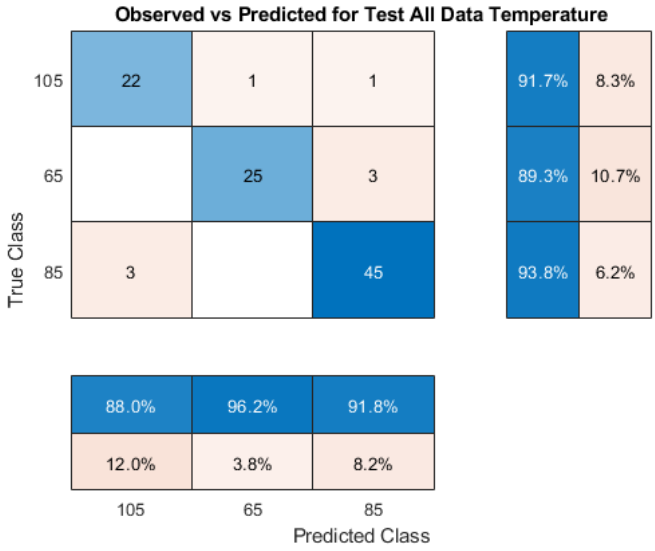

Figure 92: Data Fusion Random Forest Model Test for Temperature

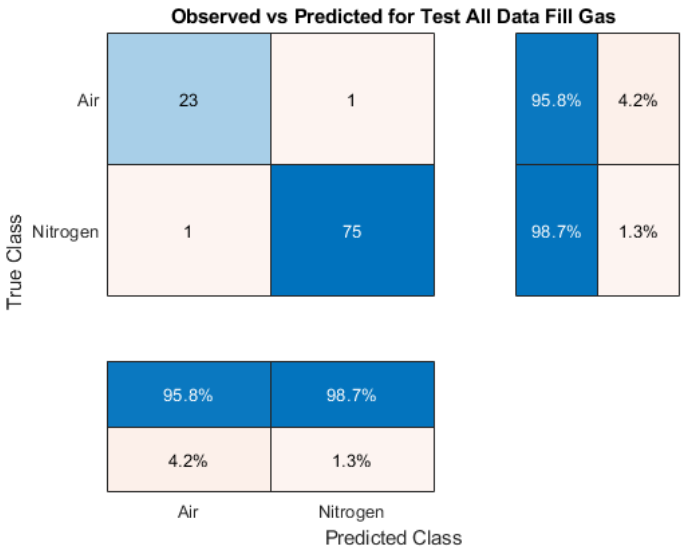

Figure 93: Data Fusion Random Forest Model Test for Fill Gas

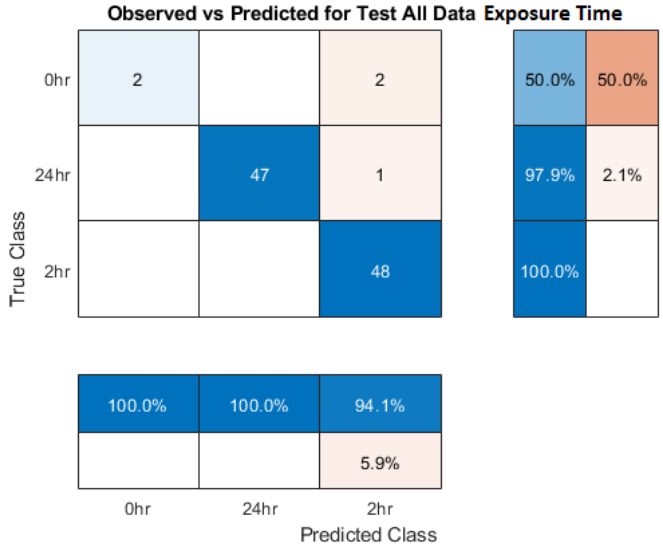

Figure 94: Data Fusion Random Forest Model Test for Exposure Time

The model works very well for most exposure times and fill gas data in its predicative capability, as seen in Figures 93 and 94. The model predicts fill gas of air and nitrogen with 98% accuracy. However, the model does have issues with for exposure time specifically with the 0hr treatments but this is possibly due to low number of

statistics. Even so it has an accuracy of 97%.

4.3.5 Random Forest Modeling Tuning

In prior parts on this experiment a standard default setting was used for each model employed; PCR, PLSR and Random Forest. With the decrease in performance by the Random Forest modeling, specifically when modeling temperature, a short study into the customization of hyper-parameters used was pursued. This was done using the same training-test split data set used in the previous analysis shown in Figures 89 and 92.

It was found that using a 1000 tree ensemble with a 90% in-bag data versus the 100 tree ensemble and 80% in-bag data yields and increase in 1% accuracy when using categorical models, and and increase of .033 or 3.3% for the R^2 when performing regression models. This tuning study it was also seen that increases in number of trees used in the ensemble was the hyper-parameter that increased performance the most.

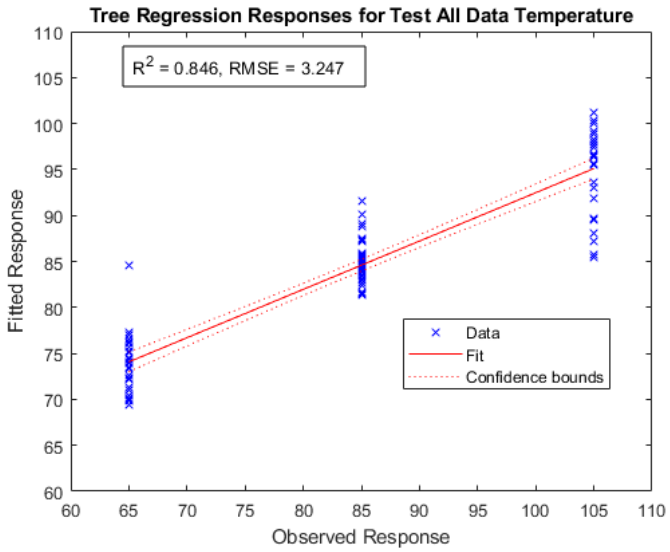

Figure 95: Data Fusion Random Forest Model Tuned Test for Temperature Regression

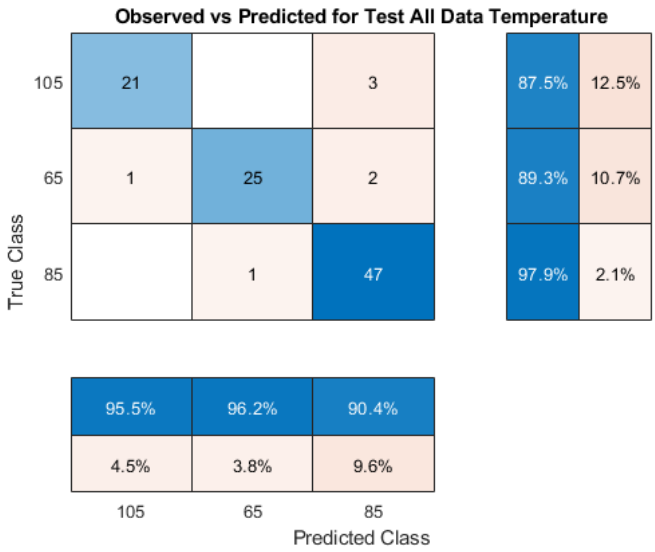

Figure 96: Data Fusion Random Forest Model Tuned Test for Temperature

While increases in performance are seen by manipulating the hyper-parameters, the overall increases were only on the order of a few percentage points. This leads the idea that more robust analysis manipulations might need to deal with characteristics such a temperature.

4.4 Temperature Modeling Study

In this experiment it is seen that temperature modeling in general was the hardest characteristic of the environmental conditions to predict. As such, further investigation into methods by which to increase predicative capabilities through more direct data manipulation are studied. For this study, the Fusion Data set previously built was parsed to create a new data set that removed variance in humidity and atmospheric composition, only allowing temperature and treatment time to vary. Thus, a data set of 120 points was made with a humidity of 35%RH and Nitrogen atmosphere, with temperatures of < 5, 35, and $70°F$, for 2 hour and 24 hour treatments. This data set was split into a training-test set and modeling was employed.

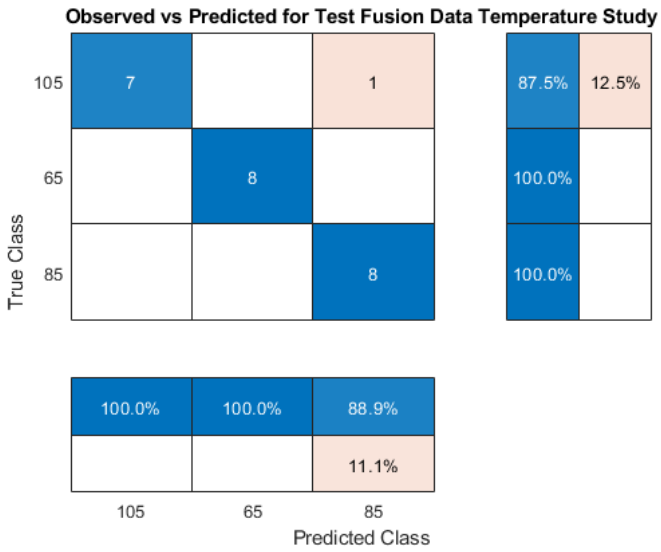

Figure 97: Data Fusion Random Forest Model Test for Temperature Study

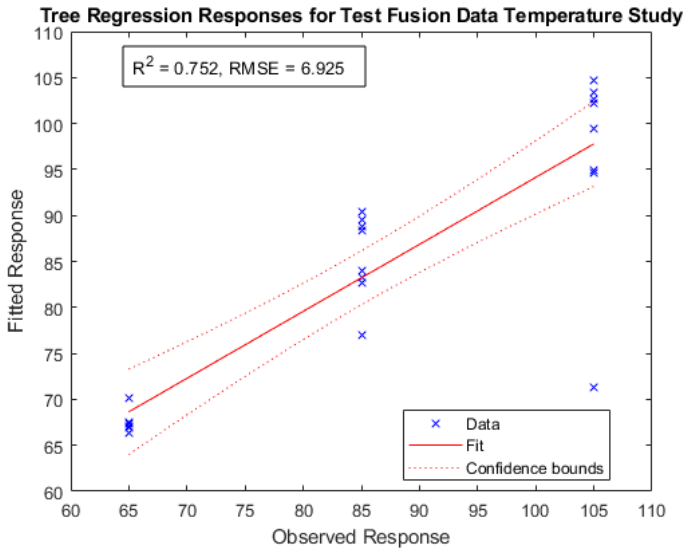

Figure 98: Data Fusion Random Forest Model Regression for Temperature Study

When looking at Random Forest models, Figures 89 and 92 show an $R^2 = .813$ for the regression model and an accuracy of 92% for the categorical model. Using the new data set for temperature, new models were built and an accuracy of 95.8% was found using categorical models, while an $R^2 = .752$ was found for regression modeling.

This show mixed results with an increase in the categorical models performance and a decrease in the regression models performance. One concern about this study is that when using a training-test set the statistics start to become very low, with only 24 data points in that test set and 96 data points in the training set. This could be indicated by the inconsistency in the modeling.

4.5 Summary

Firstly, The spectral responses in the Raman and LIBS spectra were viewed to breakdown the different trends, and it was clear that the Raman spectra could more clear define processes, due to the molecular data shown versus the atomic emissions show in LIBS spectra.

When viewing PCA for the different data breakdowns between, Raman, LIBS, and the data fusion similar trends emphasizing humidity, and treatment time as the main dictators of the separation are seen in the component space. The fused data and Raman data typically had much better separation when analysing the component space for PCA, PCR and PLSR. Compounded with this was the better separation found in each data set when compared to humidity responses versus temperature responses.

Using Table 4.5, various trends are seen. Firstly, PCR has the worst performance as a model, with a lowest $R^2 = .370$, which is expected due the methods it employs. PLSR performs the best in all cases expect for the cases of Raman data with either response variable, where the Random Forest model has the better modeling. PLSR has a max $R^2 = .995$. Also, in every case temperature models perform worse than models for humidity. This is another expected result due to the clear difference in the separations seen between humidity and temperature data, humidity being more separated.

Table 6: Data and Model Comparison

Data and Model Comparison				
Response Variable	Data Set	Model	R^3	$RMSE$
Humidity	LIBS	PCR	.784	11.015
		PLSR	.994	2.055
		Random Forest	.961	4.782
	Raman	PCR	.728	11.919
		PLSR	.983	3.482
		Random Forest	.995	1.909
	Fusion	PCR	.860	9.303
		PLSR	.995	1.796
		Random Forest	.979	3.694
Temperature	LIBS	PCR	.370	6.971
		PLSR	.977	2.144
		Random Forest	.962	2.036
	Raman	PCR	.454	7.186
		PLSR	.950	3.138
		Random Forest	.957	2.422
	Fusion	PCR	.490	7.213
		PLSR	.992	1.307
		Random Forest	.962	2.053

Another trend seen in the modeling results is that that fused data out performs the other data sets. The one exception to this is for the humidity with the PLSR model, where the Raman data performs better than the fused data.

The overall performance of the training-test split data for regression modeling was very good, with the lowest $R^2 = .492$ for LIBS data with PCR modeling temperature. The best model was PLSR with an $R^2 = 1.000$ for humidity responses using fusion data. The trend of temperature responses having worse modeling holds, as well as the fact that the fused data has the best performance of the data sets. PLSR is the highest performing model, and humidity models having higher predication capability than temperature.

By comparing the confusion charts in sections 4.1.4, 4.2.4, 4.3.4 several trends can be seen. These results are summarized in Table 4.5. In all cases the fused data

Table 7: Data and Model Comparison of Test Set

Data and Model Comparison of Test Set				
Response Variable	Data Set	Model	R^3	$RMSE$
Humidity	LIBS	PCR	.783	11.100
		PLSR	1.00	.079
		Random Forest	.863	8.458
	Raman	PCR	.741	11.793
		PLSR	.992	2.432
		Random Forest	.971	4.365
	Fusion	PCR	.862	9.222
		PLSR	1.00	.397
		Random Forest	.973	4.107
Temperature	LIBS	PCR	.492	7.273
		PLSR	1.00	.143
		Random Forest	.695	3.479
	Raman	PCR	.509	7.270
		PLSR	.961	2.797
		Random Forest	.721	4.716
	Fusion	PCR	.527	7.258
		PLSR	.999	.393
		Random Forest	.813	3.258

set provides better predictive capability when using categorical models. Also, the LIBS data performed the worst for each response variable, with Raman data being better than LIBS but worse than the fused data. Also, the differing response variable models had accuracy's from highest to lowest in the following order; humidity, fill gas, classification, exposure time, and lastly temperature. The one exception to this is in the LIBS data where the fill gas was more accurate than humidity.

Table 8: Random Forest Categorical Test Results

Random Forest Categorical Test Results		
Data Set	Response Variable	Accuracy
LIBS	Classification	.90
	Humidity	.94
	Temperature	.85
	Fill Gas	.95
	Exposure Time	.89
Raman	Classification	.92
	Humidity	.99
	Temperature	.85
	Fill Gas	.96
	Exposure Time	.91
Fusion	Classification	.97
	Humidity	1.00
	Temperature	.92
	Fill Gas	.98
	Exposure Time	.97

V. Conclusions

In this experiment, a dual Raman-LIBS experimental setup providing Raman, LIBS, and Raman-LIBS fused spectra, paired with PCR, PLSR, and Random Forest models, was able to evaluate LiH reactions with moisture, varying temperature, atmospheric compositions, for varying exposure times. It was shown that using this data collection and analysis techniques modeling can achieve predictive capabilities, for new data.

In real world applications there are instances where equipment, time, or both are limited. In general Raman spectroscopy is superior to LIBS when several minutes per sample are available and only surface data is desired, as seen through the modeling done. However, LIBS spectroscopy could have broader application ability when extended time for data collection is not available. The novel setup for conducting both LIBS and Raman using the same equipment is demonstrated with great success utilizing a single pulsed Nd:YAG laser, assorted optics, and an echelle spectrograph. The use of an echelle spectrograph, and the broad bandwidth it allows, enabled the measurement of light from 550 nm to 850 nm, which could be expanded for more wavelengths if desired. The unique dual Raman-LIBS setup allowed for efficient use of time and lab space, which is a significant advantage in labs where space and access are limited. The dual setup also minimizes potential user error when ensuring data is collected from the same sample/position for data fusion. Lastly, where commercial LIBS and Raman systems may be housed in separate facilities, this single setup minimizes potential safety concerns when transporting hazardous materials for analysis.

Using this dual Raman-LIBS setup, the data displayed in Table 4.5 shows that in most case the Data Fusion for Raman and Libs spectra provides a more accurate model, save for the case of the Random Forest model using Raman spectral data

where it had a $R^2 = .995$. While it became clear that PLSR and the Random Forest model could use any data set, Raman, LIBS, or Raman-LIBS Fusion data, PCR was seen to struggle in the case of temperature specifically where it had a $R^2 = .370$. This verifies that PCR could be used for an initial breakdown of data but, it should not be used in predictive models for this type of spectral data.

In addition, it was seen that the Random Forest model performs well when presented with a new data set. This is shown in Table 4.5. It was seen that the Raman-LIBS Fusion data performed best emphasizing the validity of a data fusion approach when modeling. This analysis also shows how the models more easily distinguish humidity while temperature is more difficult for the reactions of LiH, as was indicated by the highest $R^2 = .813$ for temperature prediction, while the highest $R^2 = .973$ for humidity. When referencing the LiH reactions chains and also the spectral lines available, this can be explained by the direct dependence on humidity as well as the spectral response when LiH is exposed to moisture. For temperature the chemical and spectral responses are more indirect or secondary, seen in an increase or decrease of certain spectral lines in the Raman and LIBS spectra.

The various characteristics modeled show how the technique employed in this experiment can be used to enhance information gained about LiH and the in-growth of lithium compounds that can occur. This work demonstrated the capability of both LIBS and Raman or a Raman-LIBS Fusion data to quantitatively characterize LiH and its reaction with environmental conditions, when combined with multivariate analysis techniques. With further development, these techniques could be applied to additional characteristics and real world data for LiH in industrial and DOD applications.

5.1 Future Work

An expansion of humidity and temperature ranges would form a more complete data set, and could be performed to study additional trends that LiH might undergo. In addition, the nitrogen and air could be augmented with other fill gases in this experiment. New fill gases such as helium and argon could be used to investigate their affects on LiH samples during treatments. In this experiment the conversion of $LiOH$ to $LiOH \cdot H_2O$ was seen with moisture exposure. One known reaction that can occur is the reverse reaction with reduced humidity and heating. Samples have been seen to show the slow conversion of the 3566 cm^{-1} spectral peak, in Raman spectroscopy, which is characteristic of $LiOH \cdot H_2O$ into a 3663 cm^{-1} peak of $LiOH$. This was shown for heating of samples at a rate of $10°C$ per min up to $100°C$, were the temperature was kept. Spectra were been recorded at different times and clearly show this reaction [14]. Performing this type of heating for samples after humidity exposures could impart additional real world characteristics and data.

While this experiment focus primarily on humidity, temperature, and some atmospheric conditions, the Raman and LIBS spectra combined with multivariate analysis should be explored for other conditions of LiH. Oxidation and reduction environments could provide useful information due the the interest of various stakeholders and their relative spheres of LiH applications.

The use of a bifurcated fiber will allow light from the same path to be sent to two separate measurement devices. The inherent disadvantage to this is the intensity of the light reaching the camera is cut in half at a minimum (assuming no losses at the fiber splice). The advantage is the ability to view spectral lines which maybe out of view of one instrument. While the echelle spectrograph has a very wide static field of view, there are monochrometers with a wider dynamic field of view.

Appendix A. Additional Raman Data Results

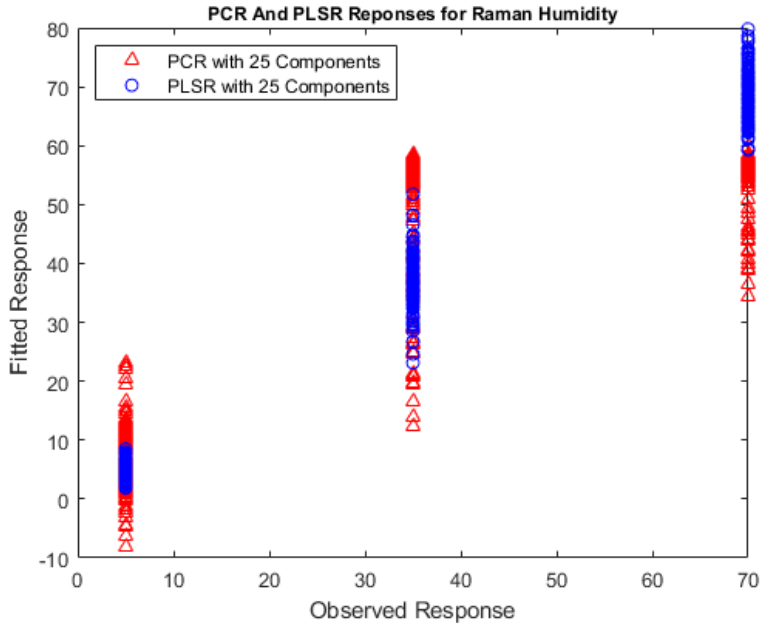

Figure 99: Raman Data PCR and PLSR Humidity Responses

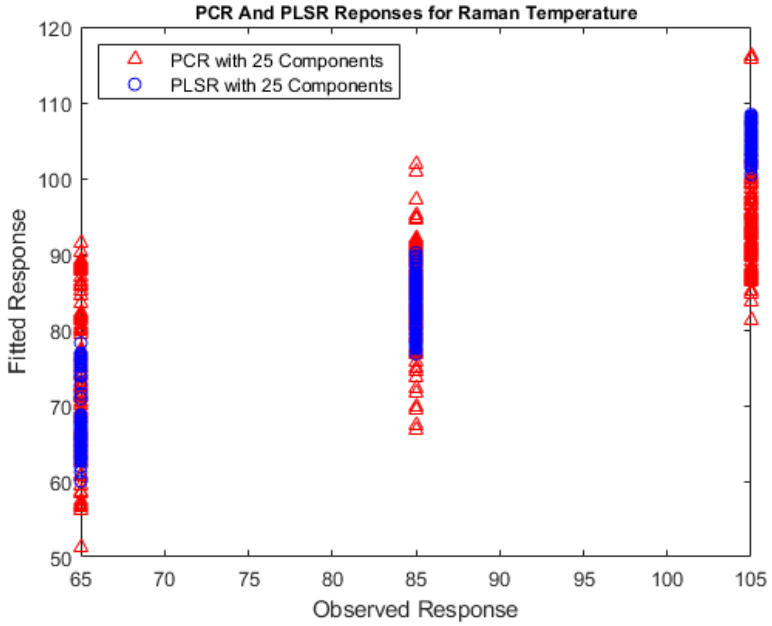

Figure 100: Raman Data PCR and PLSR Temperature Responses

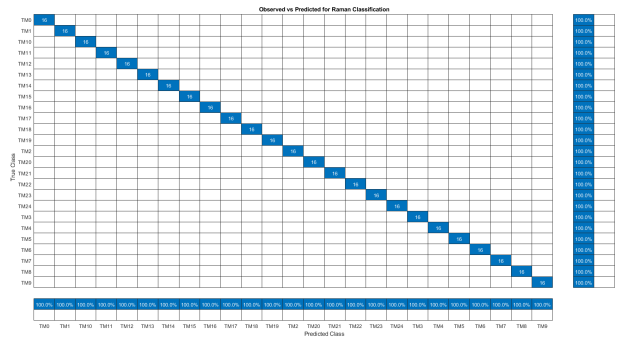

Figure 101: Raman Data Random Forest Model Classification Confusion Chart

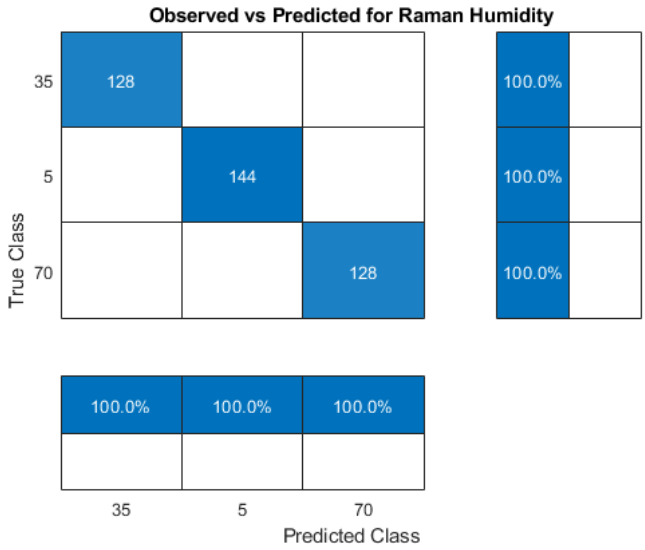

Figure 102: Raman Data Random Forest Model Humidity Confusion Chart

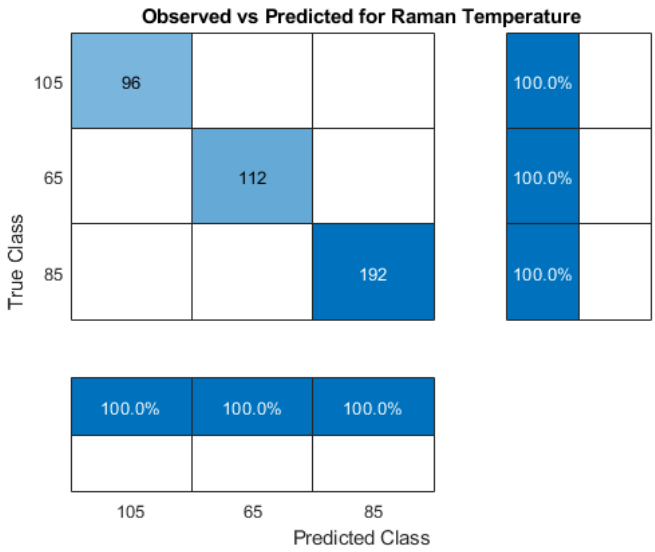

Figure 103: Raman Data Random Forest Model Temperature Confusion Chart

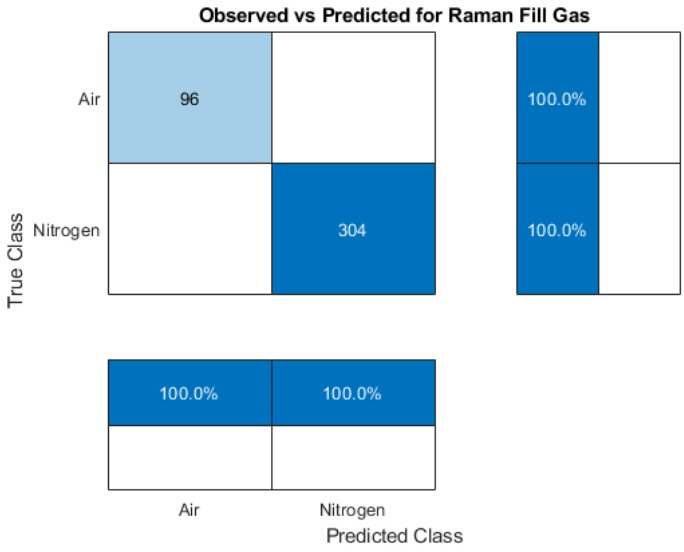

Figure 104: Raman Data Random Forest Model Fill Gas Confusion Chart

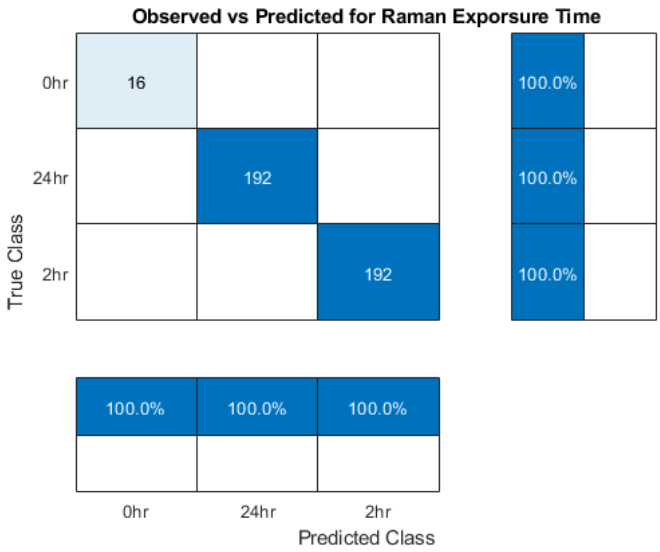

Figure 105: Raman Data Random Forest Model Exposure Time Confusion Chart

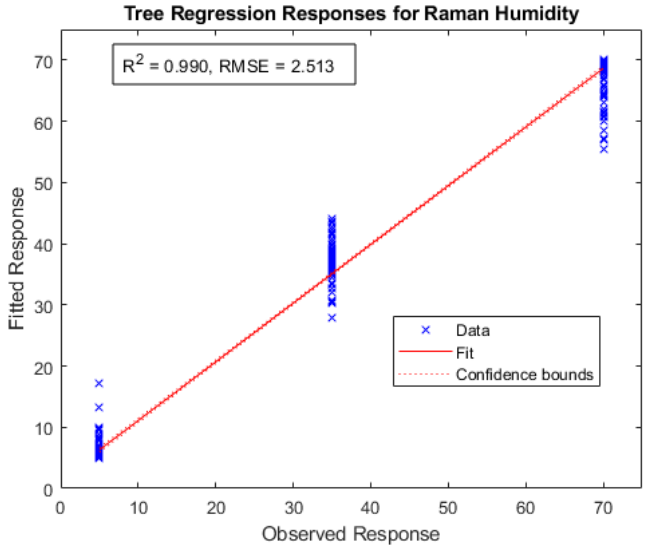

Figure 106: Raman Data Random Forest Model Regression for Humidity

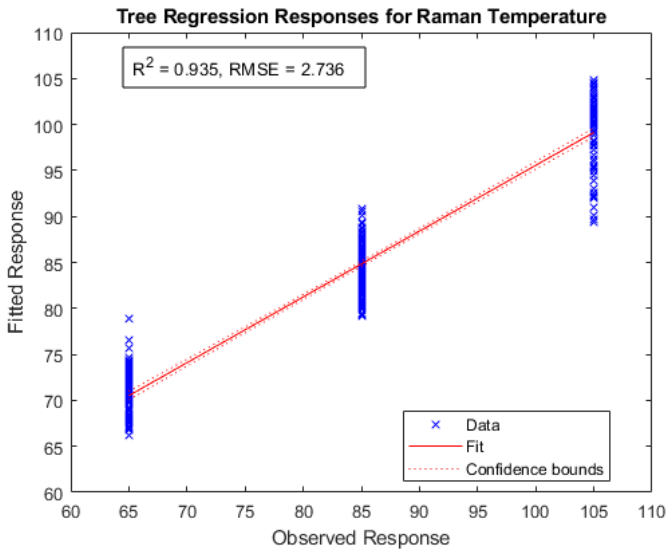

Figure 107: Raman Data Random Forest Model Regression for Temperature

Appendix B. Additional LIBS Data Results

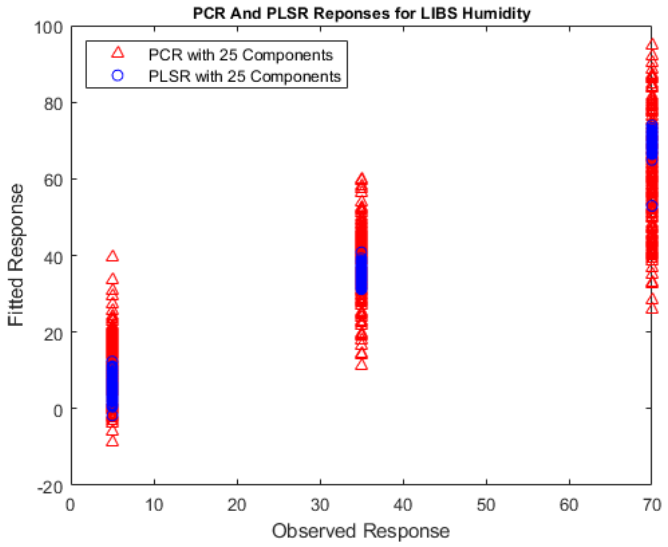

Figure 108: LIBS Data PCR and PLSR Humidity Responses

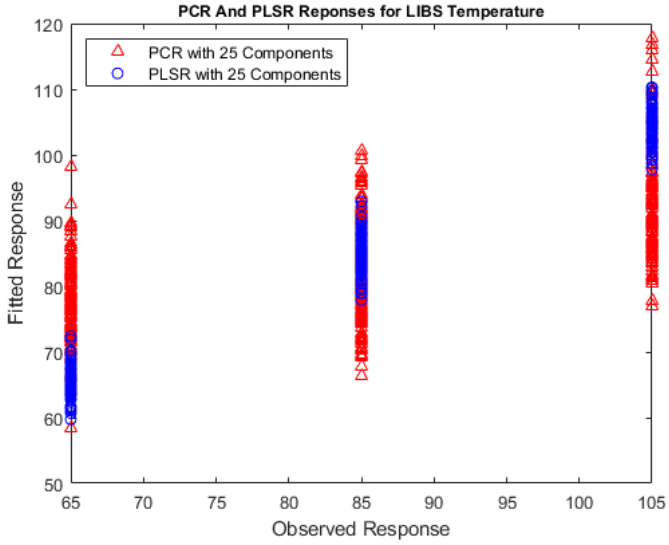

Figure 109: LIBS Data PCR and PLSR for Temperature

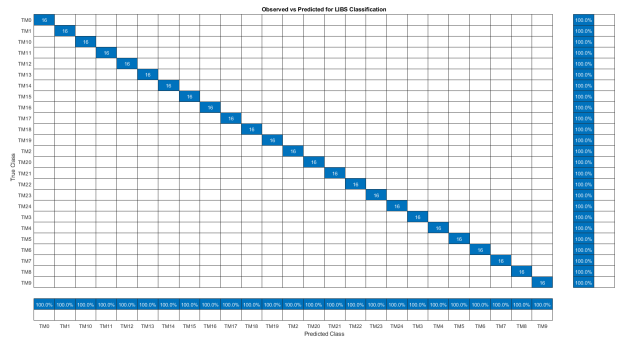

Figure 110: LIBS Data Random Forest Model Classification Confusion Chart

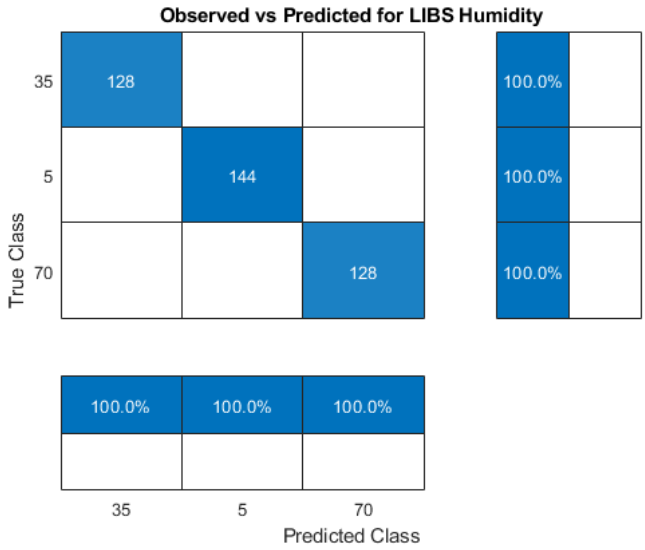

Figure 111: LIBS Data Random Forest Model Humidity Confusion Chart

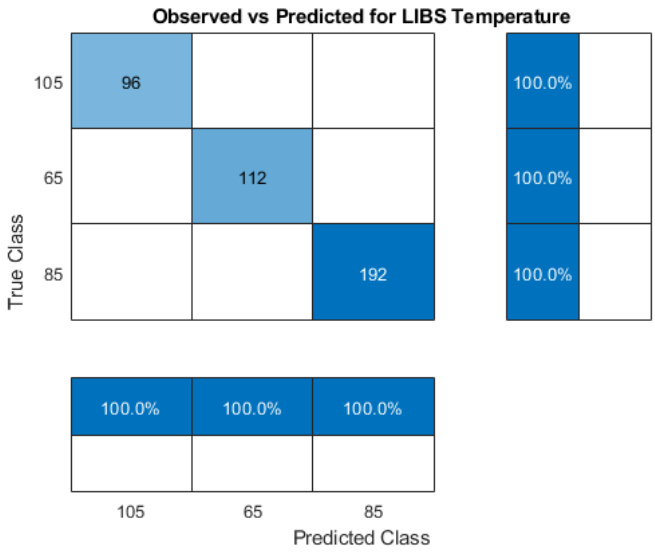

Figure 112: LIBS Data Random Forest Model Temperature Confusion Chart

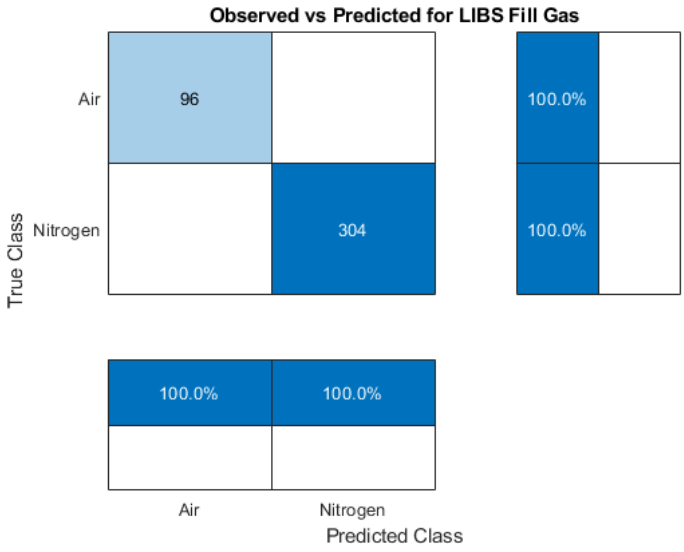

Figure 113: LIBS Data Random Forest Model Fill Gas Confusion Chart

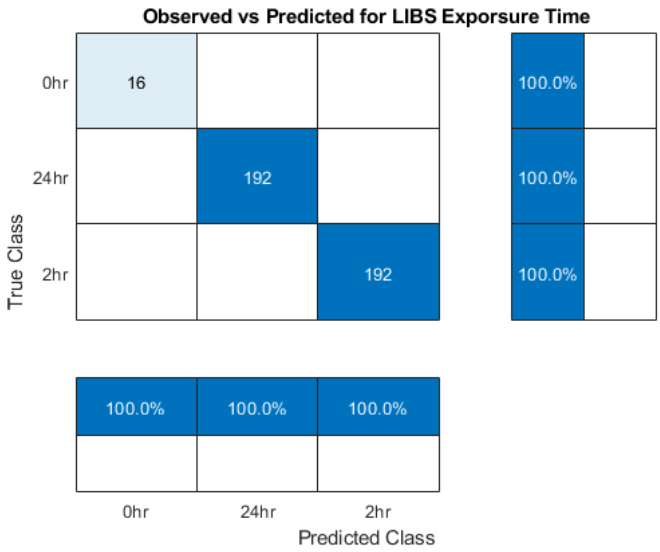

Figure 114: LIBS Data Random Forest Model Exposure Time Confusion Chart

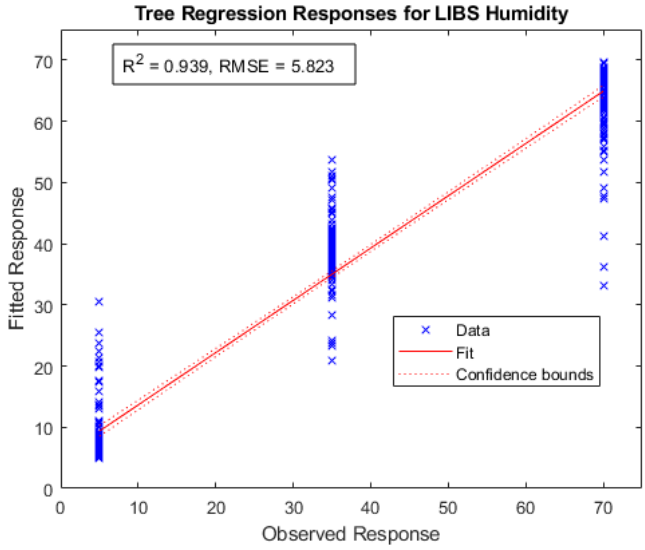

Figure 115: LIBS Data Random Forest Model Regression for Humidity

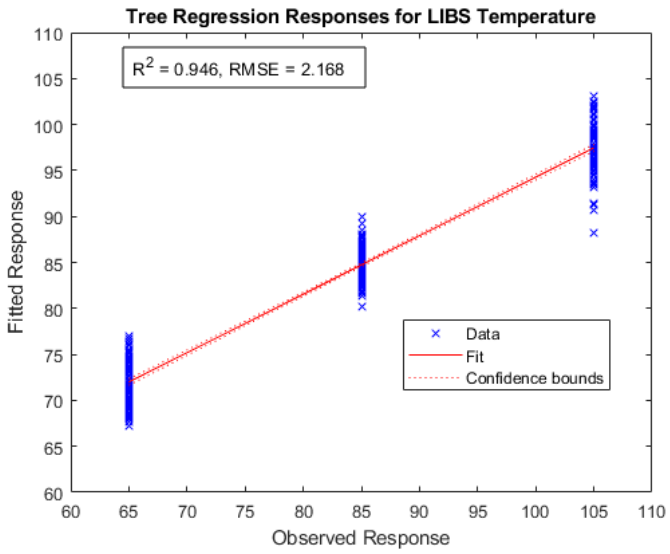

Figure 116: LIBS Data Random Forest Model Regression for Temperature

Appendix C. Additional Data Fusion Results

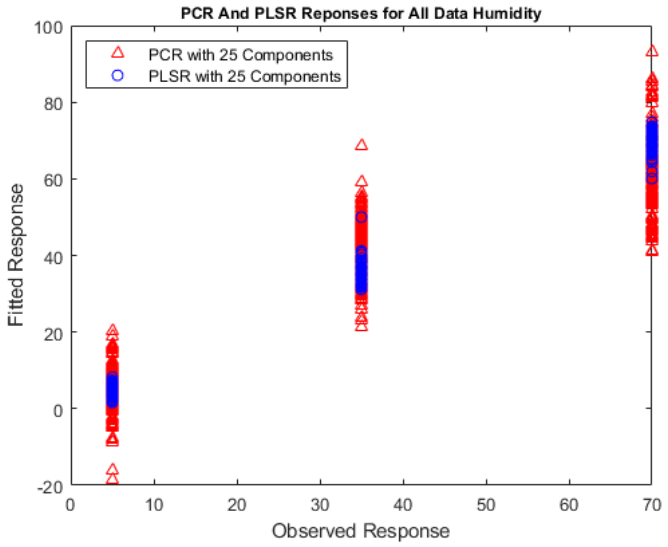

Figure 117: Data Fusion PCR and PLSR Humidity Responses

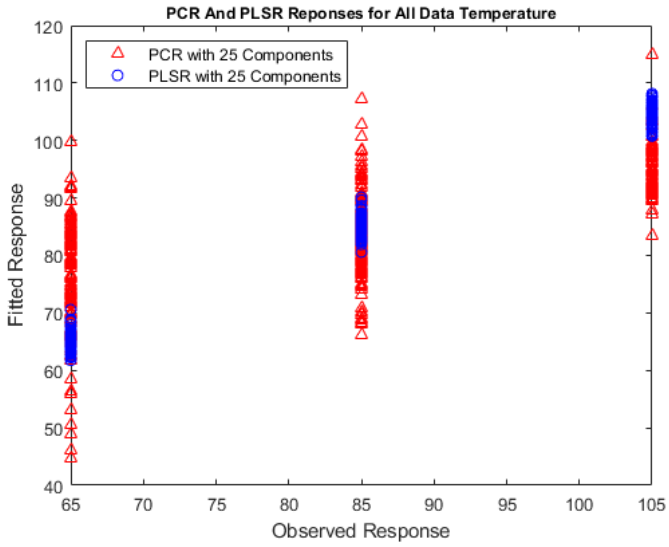

Figure 118: Data Fusion PCR and PLSR for Temperature Responses

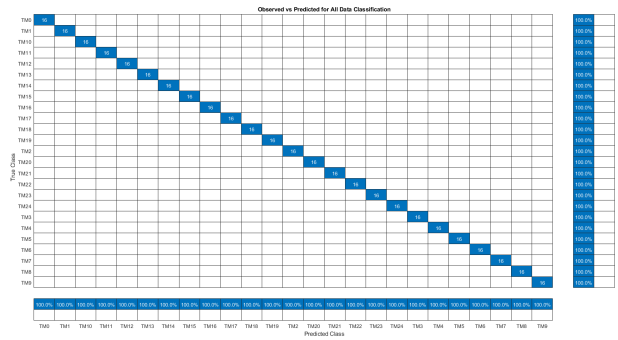

Figure 119: Data Fusion Random Forest Model Classification Confusion Chart

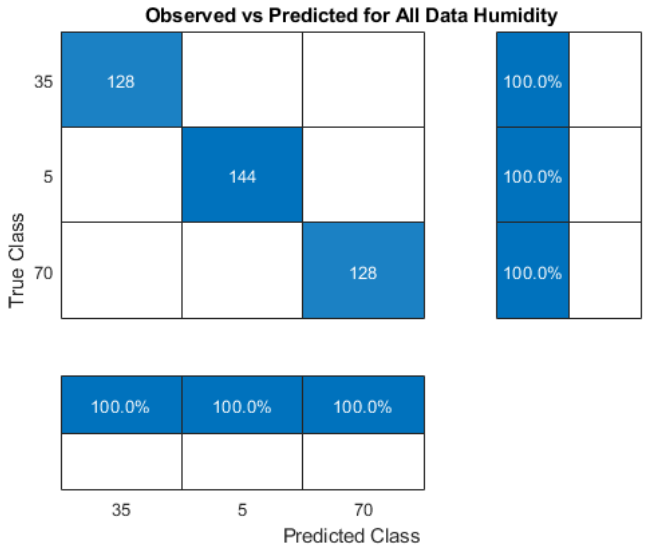

Figure 120: Data Fusion Random Forest Model Humidity Confusion Chart

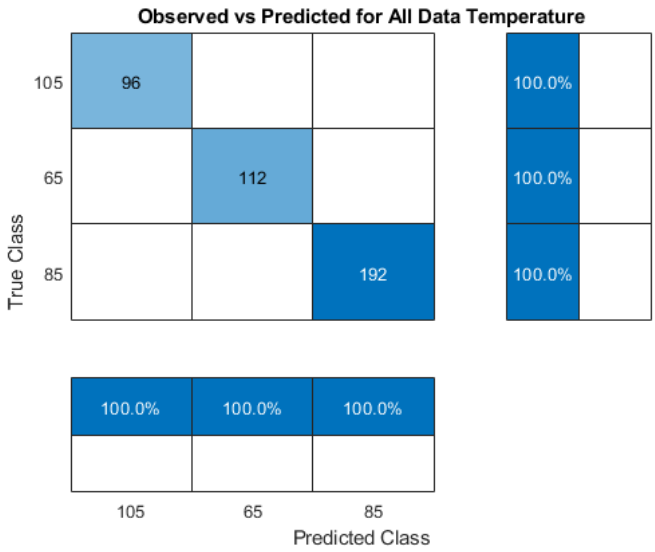

Figure 121: Data Fusion Random Forest Model Temperature Confusion Chart

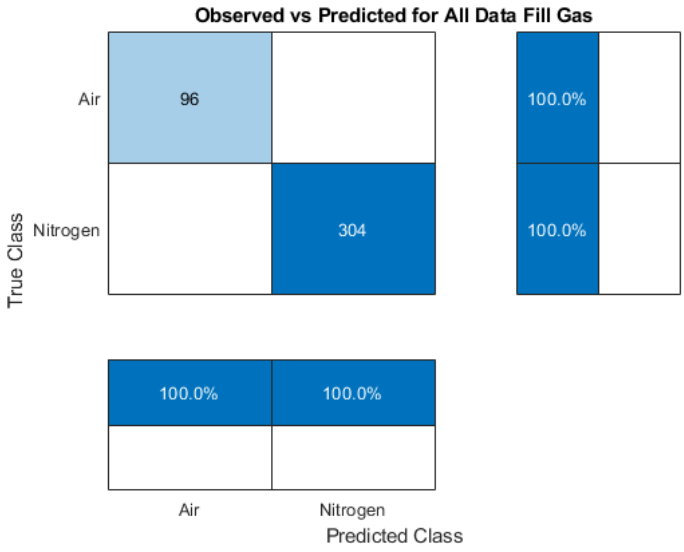

Figure 122: Data Fusion Random Forest Model Fill Gas Confusion Chart

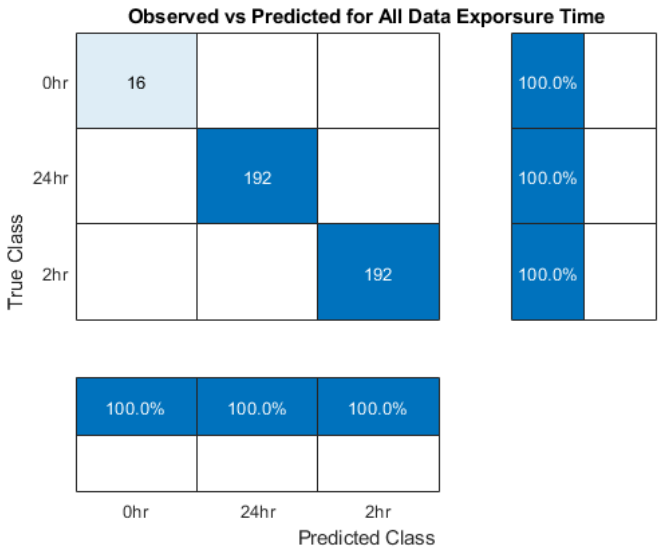

Figure 123: Data Fusion Random Forest Model Exposure Time Confusion Chart

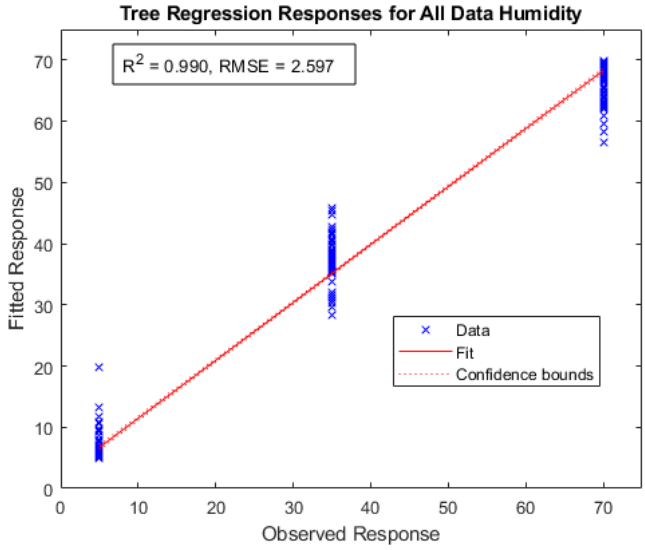

Figure 124: Data Fusion Random Forest Model Regression for Humidity

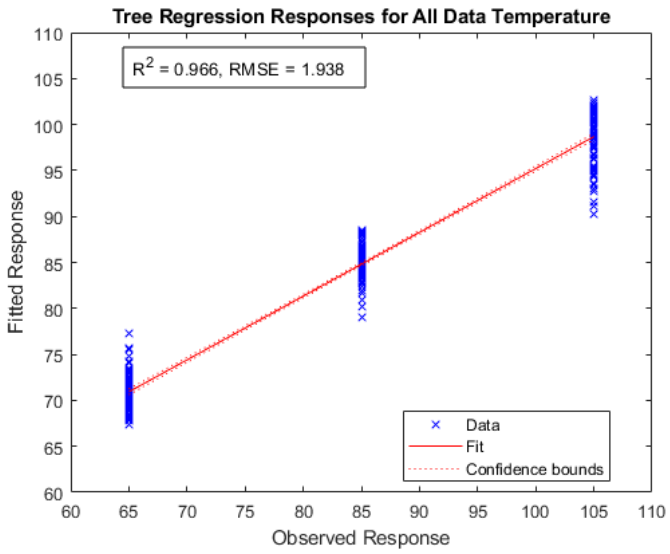

Figure 125: Data Fusion Random Forest Model Regression for Temperature

Bibliography

1. "What is raman spectroscopy?" https://www.azom.com/article.aspx?ArticleID=15797, AZO Materials., (Accessed Dec 24, 2021).

2. "What are the differences between raman and ir spectroscopy?" https://lab-training.com/2015/06/26/what-are-the-differences-between-raman-and-ir-spectroscopy/, Auriga Research Private Limited., (Accessed Dec 24, 2021).

3. "Laser-induced breakdown spectroscopy: Educational notes," https://www.princetoninstruments.com/learn/laser-induced-breakdown-spectroscopy, Teledyne Princeton Instruments., (Accessed Dec 24, 2021).

4. Chegg., https://d2vlcm61l7u1fs.cloudfront.net/media, Chegg., (Accessed Dec 24, 2021).

5. J. P. Singh and S. N. Thakur, Eds., *Laser-Induced Breakdown Spectroscopy*. Amsterdam, The Netherlands: Elsevier, 2007.

6. D. Cremers and L. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, 2nd ed. West Sussex, UK: John Wiley Sons, Ltd., 2013.

7. C. Maupoix, J. L. Houzelot, E. Sciora, G. Gaillard, S. Charton, L. Saviot, and F. Bernard, "Experimental investigation of the grain size dependence of the hydrolysis of lih powder," *Powder Technology*, vol. 208, no. 6, pp. 3415–3421, 2013.

8. V. S. Gorelik, D. Bi, Y. P. Voinov, A. I. Vodchits, B. P. Gorshunov, N. I. Yurasov, and Y. I. I, "Raman spectra of lithium compounds," *Journal of Physics: Conference Series*, vol. 918, no. 1, 2017.

9. M. H. Brooker and J. Wang, "Raman and infrared studies of lithium and cesium carbonates," *Spectrochimica Acta Part A: Molecular Spectroscopy*, vol. 48, no. 7, pp. 999–1008, 1992.

10. F. H. Welch, "Lithium hydride: A space age shielding material," *Nuclear Engineering and Design*, vol. 26, no. 3, pp. 444–460, 1974.

11. Kirk-Othmer, "Lithium and lithium compounds," in *Kirk-Othmer Encyclopedia of Chemical Technologies*, 4th ed., Kirk-Othmer, Ed. Hoboken, NJ, USA: Jon Wiley Sons, 2007, p. 450.

12. G. A. Nazri, L. A. Oumellal, A. Rougie, and T. Jean-Marie, "Lithium hydride negative electrode for rechargeable lithium batteries," United States of America Patent US7 736 805B2, June 15 2010.

13. "Lithium carbonate," Available at https://www.webmd.com/drugs/2/drug-5887-42/lithium-carbonate-oral/lithium-oral/details(accessedJan.30,2021) (2005/06/12), WebMD., accessed Dec 24, 2021.

14. S. Charton, C. Maupoix, F. Delaunay, L. Saviot, and F. Bernard, "Experimental investigation on lithium hydride hydrolysis," University of Burgundy, June 2006.

15. R. L. Smith and J. W. Miser, "Compilation of the properties of lithium hydride," Cleveland, OH, NASA Technical Memorandum X-483, Tech. Rep., January 1963.

16. M. B. Shuai, S. Xiao, Q. S. Li, M. F. Chu, and X. F. Yang, "Hydrolysis characteristics of polycrystalline lithium hydride powders and sintered bulk," *International Journal of Materials and Metallurgical Engineering*, vol. 5, no. 11, pp. 1573–1577, 2011.

17. M. Chu, D. Meng, Y. Li, M. Wang, S. Xiao, L. Luo, X. Tang, and S. Yang, "The solid reaction of lithium hydride and lithium hydroxide in lithium hydride

pellet under normal condition and the application of co_2 for long-time storage," *International Journal of Materials and Metallurgical Engineering*, vol. 447, pp. 673–676, 2018.

18. J. Guichard, F. Bouyer, E. Sciora, and F. Bernard, "Hydrolysis of lithium hydride under low relative humidity," *International Journal of Hydrogen Energy*, vol. 40, pp. 12 736–12 744, 2015.

19. S. M. Matt, J. M. Haschke, W. Meclean II, and L. N. Dinh, "Moisture corrosion of lih: A kinetic investigation by drift spectroscopy," *Journal of Physical Chemistry A*, vol. 124, no. 2, pp. 283–287, 2020.

20. A. C. Stowe and N. R. Smyrl, "Raman spectroscopy of lithium hydride corrosion: Selection of appropriate excitation wavelength to minimize fluorescence," *Vibrational Spectroscopy*, vol. 60, pp. 133–136, 2012.

21. N. Ponzio, J. Woodard, and A. C. Stowe, *Investigation of Oxidation Surface Reactions on LiH Utilizing Laser Induced Breakdown Spectroscopy (LIBS)*. Consolidated Nuclear Security, LLC, Y-12 National Security Complex and Pantex Plant, October 2014.

22. A. Sifuentes, A. C. Stowe, and N. R. Smyrl, "Determination of the role of li2o on the corrosion of lithium hydride," *Journal of Alloys and Compounds*, vol. 580, pp. S271–S273, 2013.

23. A. Anderson and F. Luty, "Raman scattering, defect luminescence, and phonon spectra of li_7h, li_6h, and li_7d crystals," *Physical Review B*, vol. 28, no. 6, pp. 3415–3421, 1983.

24. G. Weber, E. Sciora, J. Guichard, and F. Bouyer, "New insight on the lithium hydride-water vapor reaction system," *Journal of Hydrogen Energy*, vol. 43, no. 50, pp. 22 557–22 567, 2018.

25. J. L. Modisette, "Investigation of lithium hydride and magnesium as high temperature internal coolants with several skin materials," Langley Field, VA, NACA RM L58B17, Tech. Rep., May 1958.

26. C. Haertling, R. J. Hanrahan, and R. Smith, "A literature review of reactions and kinetics of lithium hydride hydrolysis," *Journal of Nuclear Materials*, vol. 349, pp. 195–233, 2006.

27. L. N. Dinh, W. McLean II, M. A. Childbach, J. D. LeMay, W. J. Siekhaur, and M. Balloch, "The nature and effects of the thermal stability of lithium hydroxide," *Journal of Nuclear Materials*, vol. 317, pp. 176–188, 2003.

28. J. W. Frazer and C. W. Schoenfelder, "Determination of lithium carbonate in lithium hydride," Available: https://babel.hathitrust.org/cgi/pt?id=mdp.39015077592841&view=1up&seq=3, Livermore, CA, UCRL-4918, Tech. Rep., 1957.

29. C. E. Holcombe, "Retardation of the reaction of lithium hydride with water vapor," Oak Ridge, TN, UC-4 Y-1835, Tech. Rep., July 1972.

30. "The effect of water vapor on the *lioh-co*$_2$ reaction.par i. dynamic isothermal system," Naval Research Lab Washington DC, October 1969.

31. M. Otto, "What is chemometrics?" in *in Chemometrics: Statistics and Computer Application in Analytical Chemistry*, 3rd ed. Jon Wiley Sons, Ltd., 2016, ch. 1, pp. 1–13.

32. D. E. Anderson, B. Ehlmann, O. Forni, S. Clegg, A. Cousin, N. Thoumas, J. Lasue, D. Delapp, R. Mcinroy, O. Gasnault, and et al, "Characterization of libs emission lines for the identification of chlorides, carbonates, and sulfates in salt/basalt mixtures for the application to msl chemcam data," *Journal of Geophysical Research: Planets*, vol. 122, no. 4, pp. 744–770, 2017.

33. T. Zhang, H. Tang, and H. Li, "Chemometrics in laser - induced breakdown spectroscopy," *Journal of Chemometrics*, vol. 32, no. 11, pp. 1–18, 2018.

34. C. B. Stipe et al, "Laser-induced breakdown spectroscopy of steel: A comparison of univariate and multivariate calibration methods," *Applied Spectroscopy*, vol. 64, no. 2, pp. 154–160, 2010.

35. C. R. Bhatt, F. Y. Yueh, and J. P. Singh, "Univariate and multivariate analyses of rare earth elements by laser-induced breakdown spectroscopy," *Applied Optics*, vol. 56, no. 8, p. 2280, 2017.

36. J. T. Stofel, "Lithium compound characterization via laser induced breakdown spectroscopy and raman spectroscopy," M.S. Thesis, Dept. Eng. Phys., Air Force Institute of Technology, Wright-Patterson AFB, OH, USA, 2021.

37. "What is raman spectroscopy," Horiba. https://www.horiba.com/en_en/raman-imaging-and-spectroscopy/, (Accessed Dec 24, 2021).

38. R. Ren, A. L. Ortiz, T. Markmaitree, W. Osborn, and L. L. Shaw, "Stability of lithium hydride in argon and air," *Journal Physical Chemistry B*, vol. 110, p. 10575, 2006.

39. "Lithium hydride. icsc 0813," Horiba. Available: https://www.ilo.org/dyn/icsc/showcard.display?p_lang=en&p_card_id=0813, International Labor Organization, April 2014, (Accessed Dec 24, 2021).

40. Deepak, "Sir chandrasekhara venkata raman: Biographical," https://www.nobelprize.org/prizes/physics/1930/raman/biographical/, The Nobel Prize Outreach., (Accessed Dec 24, 2021).

41. S. D. George and V. B. Kartha, "A hyphenated echelle libs-raman system for multi-purpose applications," *Review of Scientific Instruments*, vol. 89, no. 073108, 2018.

42. K. M. M. Shameem et al, "Echelle libs-raman system : A versatile tool for mineralogical and archaeological applications," *Talanta*, vol. 208, no. 120482, 2020.

43. D. Harvey, "Atomic emission spectroscopy," https://chem.libretexts.org/Courses/BethuneCookman_University/B-CU%3A_CH-345_Quantitative_Analysis/Book%3A_Analytical_Chemistry_2.1_(Harvey)/10%3A_Spectroscopic_Methods/10.07%3A_Atomic_Emission_Spectroscopy, LibreTexts., (Accessed Dec 24, 2021).

44. J. Wood, "Determination of lithium isotope concentration by laser induced breakdown spectroscopy using chemometrics," M.S. Thesis, Dept. Eng. Phys., Air Force Institute of Technology, Wright-Patterson AFB, OH, USA, 2020.

45. A. Kramida, K. Olsen, and Y. Ralchenko, "Nist libs database," https://physics.nist.gov/PhysRefData/ASD/LIBS/libs-form.html, National Institute of Standards and Technology., (Accessed Dec 24, 2021).

46. M. D. Luque de Castro and J. L. Luque Garcia, "Acceleration and automation of solid sample treatment," *Techniques and Instrumentation in Analytical Chemistry*, vol. 24, pp. 435–499, 2002.

47. W. Tawfik, "Why carbon lines intensities in libs are usually lower in compare to other elements like sodium?" https://www.researchgate.net/post/

Why_carbon_lines_intensities_in_LIBS_are_usually_lower_in_compare_to_other_elements_like_sodium, Cairo University., (Accessed Dec 24, 2021).

48. K. Rammelkamp et al, "Low-level libs and raman data fusion in the context of in situ mars exploration," *Journal of Raman Spectroscopy*, vol. 51, no. 9, pp. 1682–1701, 2020.

49. J. A. Manrique-Martinez et al, "Evaluation of multivariate analyses and data fusion between raman and laser-induced breakdown spectroscopy in binary mixtures and its potential for solar system exploration," *Journal of Raman Spectroscopy*, vol. 51, no. 9, pp. 1702–1717, 2020.

50. E. Gibbons, R. Leveille, and K. Berlo, "Data fusion of laser-induced breakdown and raman spectroscopies: Enhancing clay mineral identification," *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 170, p. 105905, 2020.

51. A. Smolinska et al, "Chapter 3 - general framing of low-, mid-, and high-level data fusion with examples in the life sciences," in *Data Fusion Methodology and Applications ser. Data Handling in Science and Technology*, M. Cocchi, Ed. Elsevier, 2019, vol. 31, pp. 51–79.

52. K. Dunn, "6.6 principal component regression (pcr)," https://learnche.org/pid/latent-variable-modelling/principal-components-regression, Process Improvement Using Data., (Accessed Dec 24, 2021).

53. "7.1 principal components regression (pcr)," https://online.stat.psu.edu/stat508/lesson/7/7.1, The Pennsylvania State University., (Accessed Dec 24, 2021).

54. P. Allison, "When can you safely ignore multicollinearity?" https://statisticalhorizons.com/multicollinearity, Statistical Horizons., (Accessed Dec 24, 2021).

55. A. Rao, "Rapid analysis of plutonium surrogate material via hand-held laser-induced breakdown spectroscopy," M.S. Thesis, Dept. Eng. Phys., Air Force Institute of Technology, Wright-Patterson AFB, OH, USA, 2020.

56. G. James et al., "An introduction to statistical learning with applications in r," G. Casella, S. Fiensberg, and I. Olkin, Eds. New York, NY: Springer, 2019, vol. 11, no. 4.

57. H. Abdi, "Partial least squares (pls) regression," https://personal.utdallas.edu/~herve/Abdi-PLS-pretty.pdf, University of Texas at Dallas, Dallas, TX, (Accessed Dec 24, 2021).

58. D. Lagomarsino, V. Tofani, F. Catani, and N. Casagli, *A Tool for Classification and Regression Using Random Forest Methodology: Applications to Landslide Susceptibility Mapping and Soil Thickness Modeling.* Switzerland: Springer International Publishing, January 2017.

59. R. Junjuri and M. K. Gundawar, "A low-cost libs detection system combined with chemometrics for rapid identification of plastic waste," *Waste Management*, vol. 117, pp. 48–57, 2020.

60. Y. I. Lee et al, "Interaction of an excimer-laser beam with metals. part iii: The effect of a controlled atmosphere in laser-ablated plasma emission," *Applied Spectroscopy*, vol. 46, no. 11, pp. 1597–1604, November 1992.

61. B. Lal et al., "Parametric study of pellets for elemental analysis with laser-induced breakdown spectroscopy," *Applied Optics*, vol. 43, no. 13, pp. 2792–2797, 2004.

Acronyms

LIBS Laser Induced Breakdown Spectroscopy. 28

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
25–03–2022	Master's Thesis	Sept 2020 — Mar 2022

4. TITLE AND SUBTITLE

Characterization of Environmental Conditioning of Lithium Hydride Using Spectroscopy and Machine Learning

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Pinson. Ryan E, 1st Lt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-MS-22-M-105

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Lithium compounds such as lithium hydride (LiH) and anhydrous lithium hydroxide (LiOH) have various applications in industry but are highly reactive when exposed to the atmosphere. These reactions create new molecular forms, including hydrogenous compounds and lithium carbonate (Li2CO3), which are not desirable and affect the applicability of these compounds. This work will present environmental effects such as heat, moisture, and the atmospheric conditions, storage conditions included buffer gases, pressure, and storage container materials, employing spectroscopic techniques such as Laser-Induced Breakdown Spectroscopy (LIBS). This enables retrieving critical information on the environmental effects of the Li-based compounds and the reactions that have significant implications on industrial battery technologies and nuclear security applications.

15. SUBJECT TERMS

Raman Spectroscopy, Laser Induced Breakdown Spectroscopy, Lithium, Environmental Conditioning, Machine Learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Anil K. Patnaik, AFIT/ENP
U	U	U	UU	124	19b. TELEPHONE NUMBER *(include area code)* (937) 255-6565 x4532; Anil.Patnaik@afit.edu