

Advancement of Hyperspectral Image Unmixing and Analysis: An Application in Mineral Detection and Identification

Jade Baker Preston

A dissertation for the
Doctor of Philosophy degree
presented to the faculty of the
School of Data Science
University of Virginia
on 28 February 2025

Dissertation Committee

William Basener, Advisor

Donald Brown, Chair

Jess Reia

James Lambert

Abstract

Advancement of Hyperspectral Image Unmixing and Analysis: An Application in Mineral Detection and Identification

Jade Baker Preston

This dissertation contributes to data science by addressing theoretical and practical challenges to advance hyperspectral image analysis. Hyperspectral imaging captures high dimensional spectral information at the pixel level, enabling enhanced material detection and identification. A common process in hyperspectral image analysis is spectral unmixing — the task of identifying pure materials, from an observed pixel spectrum and estimating their relative abundances.

Spectral unmixing is often framed as a regression problem, with Ordinary Least Squares (OLS) regression serving as a foundational approach. Despite its widespread use, the assumptions underlying OLS and its extensions are seldom articulated, particularly in the context of spectral unmixing. This body of work compares a variety of unmixing techniques, but also incorporates an explanation of the algorithmic assumptions and relationships between OLS and its extensions contributing to their unmixing success and failure. Through outlining the OLS assumptions, we identify alignments and misalignments with the practical demands of spectral unmixing and develop a novel technique addressing these discrepancies.

Through this exploration, the research also addresses foundational questions in data science: To what extent do algorithmic assumptions reflect real-world phenomena, and how can models balance algorithmic complexity with practical generalizability? Much of the research contributing to hyperspectral imaging involves developing or enhancing unmixing algorithms rather than evaluating them. These research questions guided the development of a comprehensive benchmarking framework. The framework evaluates techniques using metrics such as root mean squared error (RMSE), computation time, model size, percent detection, and average precision of the top-K results. To date, no study has integrated this breadth of unmixing techniques and evaluation metrics into a single framework.

Beyond evaluation metrics, we contribute to the theoretical foundation of unmixing by examining the physical-chemical phenomena contributing to material identification. We develop a spectral taxonomy classifying minerals based on their molecular design structure, providing insights into spectral material patterns and detection strengths of various techniques. Our analysis reveals two core findings: 1) technique success depends on the primary objective of the study and 2) the minerals included in the mixture model were either the target mineral (successful detection), a substitution from the same mineral category as the target, or a mineral with similar spectral pattern features (failed detection).

Acknowledgments

I would like to acknowledge my husband, parents, and siblings for their love and faith in me. Additionally, I'd like to acknowledge my advisor for his continued support in this academic journey. I am forever grateful to my husband Stephen Preston, my parents James and Paula Baker, my grandparents Howard and Melva Johnson, James and Frances Baker, my siblings Jasmine and Ryan Mossbarger, Rashad and Nikki Tomlin and my advisor Bill Basener. To God be the glory!

Common Acronym and Abbreviation List

1. ASD - Analytical Spectral Device
2. AVIRIS - Airborne Visible/Infrared Imaging Spectrometer
3. BFS - Breadth-First Search
4. BMA - Bayesian Model Averaging
5. DFS - Depth-First Search
6. FOV - Field of View
7. GSD - Ground Sample Distance
8. JPL - Jet Propulsion Laboratory
9. LASSO - Least Absolute Shrinkage and Selection Operator
10. LMM - Linear Mixture Model
11. MINLP - Mixed Integer Nonlinear Program
12. MSE - Mean Squared Error
13. NLMM - Nonlinear Mixture Model
14. nm - Nanometers
15. NNLS - Non-negative Least Squares
16. OLS - Ordinary Least Squares
17. RMSE - Root Mean Squared Error
18. SNR - Signal-to-Noise Ratio
19. SWIR - Shortwave Infrared
20. USGS - United States Geological Survey
21. VNIR - Visible and Near-Infrared

Contents

Committee	ii
Abstract	iii
Acknowledgments	iv
Acronym and Abbreviation List	v
List of Figures	viii
List of Tables	x
1 Introduction	1
1.1 Importance of Hyperspectral Imaging	1
1.1.1 Imaging Platforms	2
1.1.2 Hyperspectral Imaging Foundations	3
1.1.3 Light Terminology, Attributes, and Measurements	5
1.1.4 Data Collection and Sampling	7
1.1.5 Spectral and Spatial Resolution Trade-off	9
1.2 Mixed Pixel Phenomenology	11
1.3 Data Sources	13
1.4 Spectral Unmixing	14
1.4.1 Linear Mixture Model	15
1.4.2 Nonlinear Mixture Model	17
1.5 OLS for Unmixing	19
1.5.1 OLS Assumptions	21
1.5.2 Practical LMM Constraints	24
1.5.3 Challenge with OLS Unmixing	25
1.6 Alternative Unmixing Techniques	27
1.6.1 Regularization and Sparse Approaches	27
1.6.2 Iterative Approaches	29
1.6.3 Mixed Integer Nonlinear Program (MINLP)	32
1.6.4 Bayesian Model Averaging (BMA)	33
1.7 Overview to Dissertation	34
1.7.1 Research Gaps	34
1.7.2 Contribution Outline	34

2	Theoretical and Practical Progress in Spectral Unmixing from a Sparse Perspective	36
2.1	Introduction	36
2.2	Methods	36
2.2.1	Data Collection	36
2.2.2	Unmixing Algorithms	37
2.3	Results	38
2.4	Discussion	41
3	Unmixing with Large Spectral Libraries for Identifying Soil and Mineral Chemistry	43
3.1	Introduction	43
3.2	Unmixing Techniques	43
3.3	Physical-chemical Taxonomy	44
3.4	Results	45
3.4.1	Unmixing Performance	45
3.4.2	Feature Taxonomy	52
3.5	Discussion	52
3.6	Conclusion	53
4	Rethinking Unmixing: A Comprehensive Evaluation of Algorithms	54
4.1	Introduction	54
4.2	Experimental Design	54
4.3	Statistical Analysis	57
4.3.1	LASSO Regression	57
4.3.2	DFS Strategy	57
4.3.3	MINLP	57
4.3.4	HySUDeB	59
4.4	Technique Evaluation	60
4.5	Results	60
4.6	Discussion	64
5	Conclusion and Future Work	66
5.1	Summary of Contributions	66
5.2	Future Work	68
5.3	Concluding Remarks	69
	References	70
	A Appendix	82

List of Figures

1.1	Display of a conceptual hyperspectral image and the associated material spectra from pixels in the image.	1
1.2	Illustration of remotely sensed spectral data using aircraft imagery.	3
1.3	Simplified visualization showing the two components of an electromagnetic wave: the electric and magnetic fields. The oscillations of these fields are perpendicular to each other and to the direction of wave propagation [26]–[28].	6
1.4	Illustration of the electromagnetic spectrum showcasing various types of electromagnetic radiation, from gamma rays to radio waves. Accompanying images (beneath each category name) visually represent the relative sizes of their wavelengths, highlighting the diverse range of frequencies and energies that characterize different regions of the spectrum.	7
1.5	Display of the reflectance values of specific wavelengths corresponding to blue, green, and red color bands on a conceptual hyperspectral image.	8
1.6	Depiction of the spectral signature from a soil pixel on a conceptual hyperspectral image. The plot illustrates the percent reflectance of light across a wavelength range from 400 to 2500 nm.	9
1.7	Presentation of two linear mixture examples: car on the road illustration (1.7a) and adjacent mineral composition (1.7b)	12
1.8	Example of two types of mixture scenarios. Figure 1.8a is an example of a linear mixture of two materials. Figure 1.8b is a potential nonlinear mixture example where the components in the image are intimately mixed.	13
1.9	Display of granular mineral compositions: alunite, kaolinite, and a mixture of the two minerals (from left to right). Underneath each image is their associated pixel spectrum.	18
1.10	Display of OLS unmixing results. The bold blue spectrum represents the observed pixel. The dashed spectrum represents the inferred model. The non-bold colored spectra represent the material included in the inferred model.	26
1.11	Comparison of search optimization strategies: Breadth-first search (a) and Depth-first search (b).	30
2.1	Display of the region of selected pixels using “Hyperspectralpy” [44]. The shaded orange region indicates the selected alunite pixels and the shaded purple region indicates the kaolinite pixels.	37

2.2	Comparison of the unmixing results based on RMSE, runtime, and model size. The left plot displays the results for alunite. The right plot displays the results for kaolinite.	40
3.1	Display of the mineral types sulfate, hydroxide, inosilicate, phyllosilicate, sorosilicate, and cyclosilicate. Figure 3.1a shows the target mineral alunite highlighted in purple. Figure 3.1d shows the target mineral kaolinite highlighted in blue.	46
3.2	Alunite Table and Plot. Displays the inferred spectrum using LASSO regression with the most common minerals and their average abundances. Rows containing alunite are highlighted in yellow.	47
3.3	Kaolinite Table and Plot. Displays the inferred spectrum using LASSO regression with the most common minerals and their average abundances. Rows containing kaolinite are highlighted in cyan.	48
3.4	Alunite Table and Plot. Displays the inferred spectrum using ElasticNet with the most common minerals and their average abundances. Rows containing alunite are highlighted in yellow.	49
3.5	Kaolinite Table and Plot. Displays the inferred spectrum using ElasticNet with the most common minerals and their average abundances. Rows containing kaolinite are highlighted in cyan.	50
4.1	Display of the selected ROIs for three target minerals: alunite, kaolinite, and montmorillonite. The top row of plots show the regions where the target mineral shows higher abundance and the bottom row depicts the actual selected pixels from the locations with higher presence of the target mineral.	55
4.2	Comparison of the mineral types sulfate, tectosilicate, nesosilicate, and phyllosilicate. The purple spectra represents alunite and the dotted blue spectra are minerals that were often incorporated in the models.	56
4.3	Comparison of the performance of each unmixing technique based on error and computation time.	61
4.4	Display of the count of the included minerals across all the implemented approaches.	63
4.5	Display of the count of the mineral from the library that were included in the models for the alunite ROI.	64

List of Tables

1.1	Categorization of implemented techniques by unmixing approach	35
2.1	Overview of the unmixing techniques with their associated RMSE loss functions.	39
2.2	Displays the comprehensive results of the unmixing techniques for the alunite and kaolinite ROIs.	39
3.1	Presentation of alpha parameters for each unmixing technique. The parameter α controls the regularization strength for techniques that use regularization, while ‘-’ indicates methods where α is not applicable.	44
3.2	Comparison of unmixing techniques with metrics including RMSE mean, model size, runtime, and detection rate for two datasets.	45
4.1	Detection percentages for target minerals (alunite, kaolinite, and montmorillonite) across four techniques: Lasso, DFS, MINLP, and HySudeB.	61
4.2	Average mean precision at k for minerals (alunite, kaolinite, and montmorillonite) across four techniques: Lasso, DFS, MINLP, and HySudeB.	62
A.1	Complete comparison of unmixing techniques for each of the three ROIs. . .	83

1 Introduction

This introduction provides a thorough foundation of hyperspectral imaging concepts.

1.1 Importance of Hyperspectral Imaging

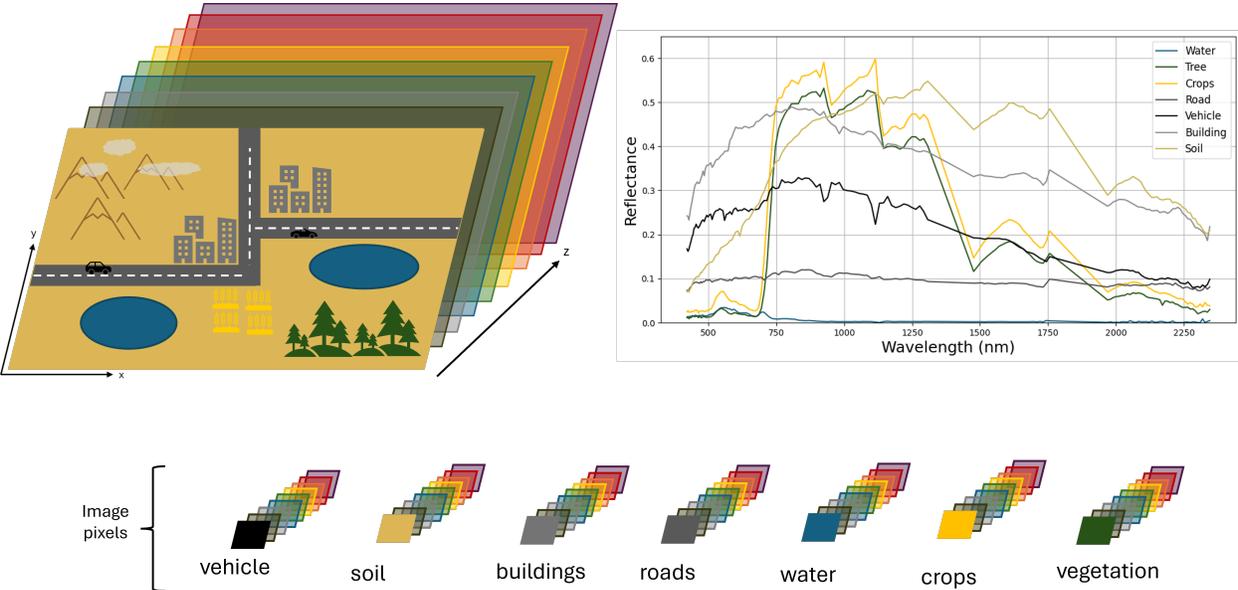


Figure 1.1: Display of a conceptual hyperspectral image and the associated material spectra from pixels in the image.

Everything on the earth’s surface, natural and man-made, has a unique spectral signature or spectrum that is characteristic to its material properties. A spectral signature is a substance’s distinctive pattern of reflected or emitted electromagnetic radiance arriving at a sensor’s field of vision (FOV), and measured over a continuous range of wavelengths [1],

[2]. The features defining each spectral signature indicate the molecular structure, chemical properties, physical attributes, and environmental conditions associated with the substance. Because of this key environmental phenomenon, hyperspectral imaging has made great advances in the fields of Spectroscopy and Remote Sensing over the past three decades.

Hyperspectral imaging is an imaging technology that captures the electromagnetic radiation across a wide range of wavelengths, enabling the detailed characterization of materials based on their spectral signatures. Figure 1.1 provides a conceptual display of hyperspectral imaging, illustrating the spectral signatures associated with individual substances. Each spectral signature in the figure corresponds to a pixel representing a single substance (though, in practice, pixels often contain multiple substances). Natural material — i.e. vegetation, water, soil, and atmosphere — as well as man-made material — i.e. buildings, roads, and vehicles — follow different spectral patterns. The high resolution characteristic to hyperspectral imaging allows for increased precision of material identification and delineation. Ultimately, hyperspectral images have played a crucial role in mineral detection, agricultural and environmental monitoring, pollution surveillance, medicine, military applications, and water purity assessment [3]–[5].

1.1.1 Imaging Platforms

Spectroscopy can be characterized as the analysis of electromagnetic radiation mapped as a function of wavelength. This mapping — also known as spectral signatures — is a measurement of the interaction between electromagnetic radiation and matter [6], [7]. Remote Sensing is the study of acquiring and understanding spectral information at a distance or without direct contact [8], [9]. Spectral data can be obtained from sensors embedded in a multitude of platforms to include remote sensing instruments — satellites, aircrafts, drones, balloons, or space shuttles — as well as direct sensing instruments — laboratory, car-mounted, or hand-held spectrometers [10]–[13]. Figure 1.2 displays a theoretical application of remotely sensed data accessed via aircraft imagery. Two of the most common tools used to acquire spectral

data — and the instruments focused on in this study — are via handheld spectrometers and airborne observation.

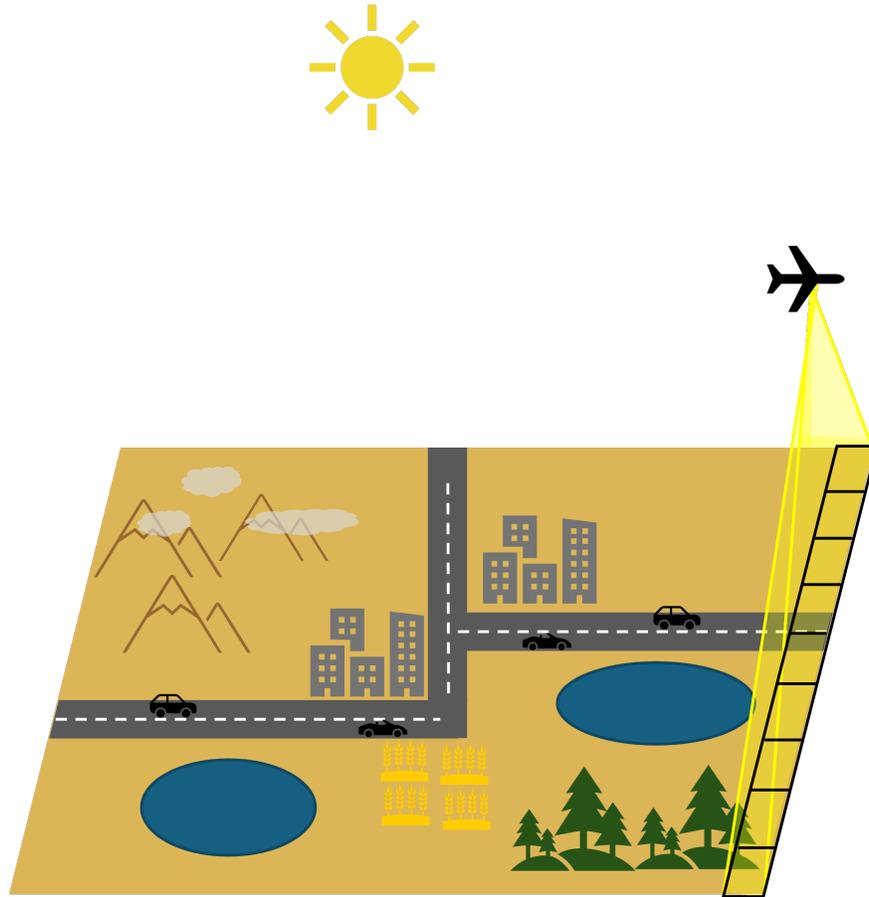


Figure 1.2: Illustration of remotely sensed spectral data using aircraft imagery.

1.1.2 Hyperspectral Imaging Foundations

Remote Sensing Developments

Hyperspectral imaging is considered a remarkable advancement evolving from the broader field of remote sensing. Though remote sensing — coined in the 1960s by scientist Evelyn Pruitt — has a long history arguably tracing back to ancient times, hyperspectral imagery has a distinct origin stemming from the inadequacies of multispectral imagery [14]–[16]. Charles Cohen credits the human eye as the start of remote sensing and notes certain revolutionaries

such as Galileo and Newton as milestone contributors [15]. Campbell et al. begin their remote sensing timeline in the 1800s with Sir William Herschel’s infrared discovery [14]. Arthur Cracknell attributes substantially later contributions such as Neubronner’s homing pigeons in 1907, polar orbiting beginning in the 1960s and weather satellites in the 1970s [17]. Conversely, hyperspectral imaging is a relatively modern innovation, which emerged to address the limitations of multispectral systems such as Landsat Multispectral Scanner [18].

Multispectral Developments

David Landgrebe authored numerous papers discussing multispectral analysis and later its implications to hyperspectral imagery. He attributed the development of multispectral analysis and satellite observation to four key historical factors: the creation of the digital computer, successes in pattern recognition, the race for space superiority, and the need for global resource management [19]–[22]. Space-based technology and eventually multispectral imagery was an efficient economic response to obtaining and processing larger amounts of the earth’s data. Satellites provided the ability to procure increased regional information on varying targets sizes (i.e crops), but resulted in a costly resolution payload [21], [22]. Thus the researchers began to glean information from electromagnetic “fingerprints” produced by the interaction of light energy and matter, also known as spectral data [18].

Hyperspectral Developments

Shen-En Qian explains the progression to hyperspectral image analysis resulted from the inability to discriminate mineral components or soil mixtures obtained from multispectral imagery. Researchers recognized the enhanced resolution capabilities of portable or laboratory spectrometers compared to satellite sensors and subsequently sought to improve remote sensing platforms. Qian also provides a detailed overview of air and space borne hyperspectral imagers. The first airborne hyperspectral imagers were developed in the 1980s led by the United States Airborne Imaging Spectrometer (AIS) and Airborne Visible/Infrared Imaging

Spectrometer (AVIRIS) and Canadian Fluorescence Line Imager (FLI), Compact Airborne Spectrographic Imager (CASI) and Shortwave-infrared Full Spectrum Imager (SFSI). Later in the 1990s, hyperspectral imagers aboard satellites were produced. At the time of Qian’s article (2021), 25 spaceborne hyperspectral imagers existed, with five more scheduled to be launched in this decade [18]. Since Qian’s article, Environmental Mapping and Analysis Program (EnMAP) was launched via SpaceX Falcon 9 rocket in April 2022, Moons And Jupiter Imaging Spectrometer (MAJIS) was launched via JUperiter ICy moons Explorer in April 2023 and Ocean Color Instrument (OCI) aboard Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) spacecraft was recently deployed in February 2024 [23]–[25].

1.1.3 Light Terminology, Attributes, and Measurements

Electromagnetic radiation — often referred to as light — consists of energy emitted, absorbed, or reflected by substances in the form of tiny particles called photons. This energy propagates through space as electromagnetic waves [26], [27]. Photons represent the smallest discrete units of electromagnetic energy [28], [29], and their propagation refers to the wave movement and transfer of energy from a source through space. Electromagnetic waves are defined by perpendicular, oscillating electric, and magnetic fields [26], [30]. Figure 1.3 illustrates the oscillations of the electric field (in red) and the magnetic field (in blue). The black arrow indicates the direction of wave propagation, also known as energy flow or photon movement [26], depicting the relationship between these fields in electromagnetic waves.

Electromagnetic waves can be characterized by their frequency and wavelength, which together determine their position on the electromagnetic spectrum. The electromagnetic spectrum includes seven main categories: gamma rays, x-rays, ultraviolet radiation, visible light, infrared radiation, microwaves, and radio waves [26]–[28]. Figure 1.4 provides a visualization of classifications within the electromagnetic spectrum. The left side of the spectrum represents the shortest wavelengths, which correspond to the highest frequencies. Wavelengths on this end (i.e. ultraviolet and x-rays) can pose greater harm to humans due to

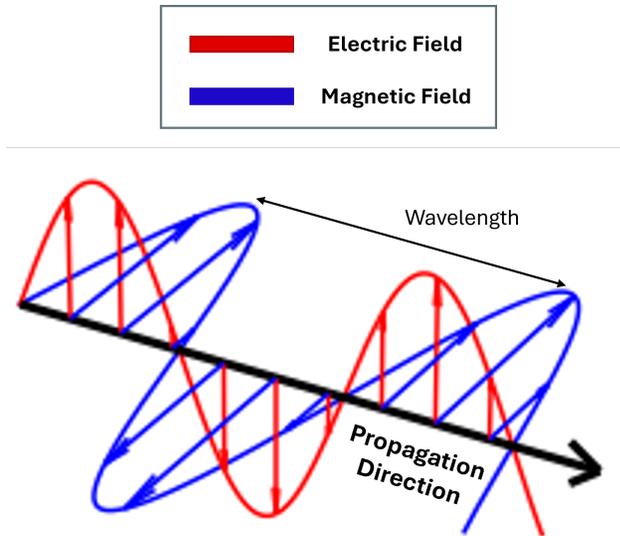


Figure 1.3: Simplified visualization showing the two components of an electromagnetic wave: the electric and magnetic fields. The oscillations of these fields are perpendicular to each other and to the direction of wave propagation [26]–[28].

their high frequency [30]. The visible spectrum — which lies between ultraviolet and infrared wavelengths — represents the portion of wavelengths that are detectable to the human eye. These visible wavelengths range roughly from 400 to 700 nanometers (nm). The right side of the spectrum encompasses the longest wavelengths, which correspond to lower frequencies of radiation [26], [28].

Radiance — measured in watts per square meter per steradians ($\text{W}/\text{m}^2/\text{steradian}$) — is the quantified measurement of electromagnetic radiation reflected or emitted from a surface in a specific direction. Hyperspectral imaging analysis focuses primarily on radiance reflected off surfaces. Consequently, radiance is often converted to reflectance, a unit-less quantity ranging from zero to one. This converted measurement represents the proportion of radiation or light reflected by the surface and typically expressed as a percentage [14], [31]. Figure 1.1 illustrates in the line graph the percent reflectance measurements for each spectral signature at their corresponding wavelengths.

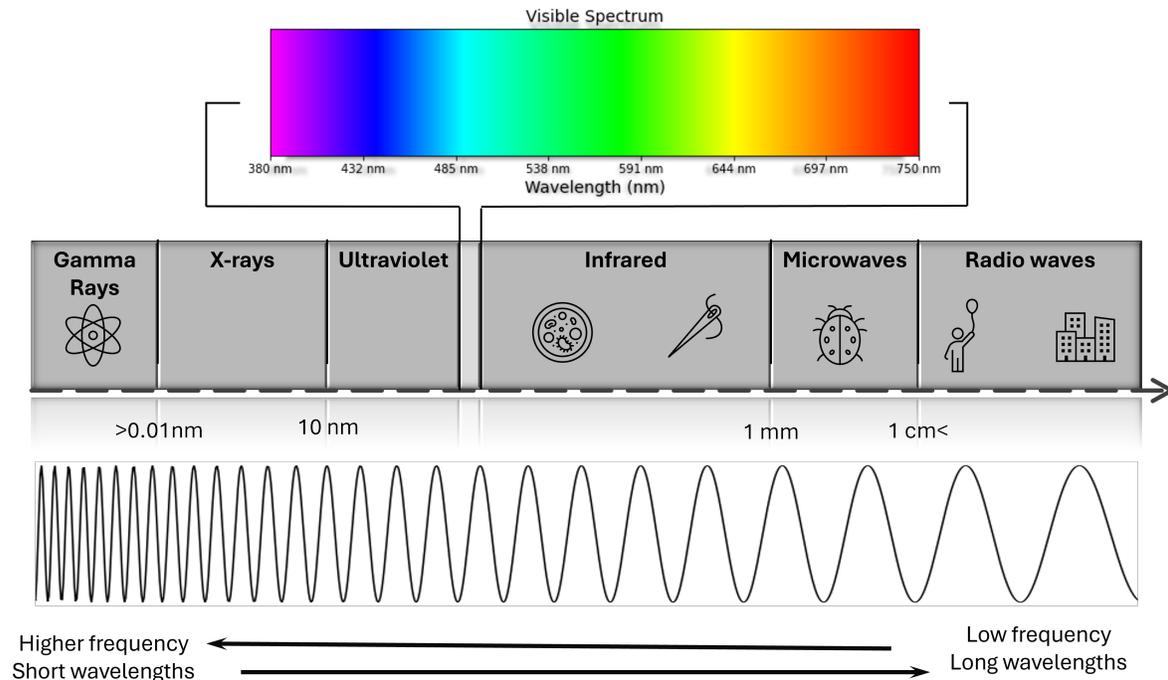


Figure 1.4: Illustration of the electromagnetic spectrum showcasing various types of electromagnetic radiation, from gamma rays to radio waves. Accompanying images (beneath each category name) visually represent the relative sizes of their wavelengths, highlighting the diverse range of frequencies and energies that characterize different regions of the spectrum.

1.1.4 Data Collection and Sampling

Researchers collecting spectral data implement four main sampling processes: spectral, spatial, temporal, and radiometric [28], [31]. Each process addresses a different aspect of data collection — spectral sampling measures reflectance across various wavelengths; spatial sampling captures spectral information from different locations or pixels; temporal sampling accounts for changes over time; and radiometric sampling ensures the accuracy and consistency of reflectance measurements [7], [31]. The two methods focused on in this study are spectral and spatial sampling.

A simplified example of spectral sampling can be described using the visible spectrum of light which encompasses wavelengths corresponding to the spectral bands of violet, blue, green, yellow, orange, and red. Figure 1.5 displays discrete points representing the percent

reflectance at three contiguous narrow spectral bands within the blue, green, and red visible range. These spectral bands correspond to specific ranges of wavelengths on the electromagnetic spectrum. Due to the individual properties of the materials in the conceptual image, objects appear dark or have lower percent reflectance regardless of the spectral band. In practice, procured spectral data are measured by the reflectance percentage over a large number of spectral bands [1], [31] — not just three or seven bands. The amount of spectral bands can vary. Researchers in the Hyperspectral field of study focus on spectral data mapped over hundreds of spectral bands typically spanning between 0.4-2.5 micrometers (μm) or 400-2500 nm. This region of bands includes the visible, near-infrared (VNIR), and most of the shortwave infrared (SWIR) light spectrum [3].

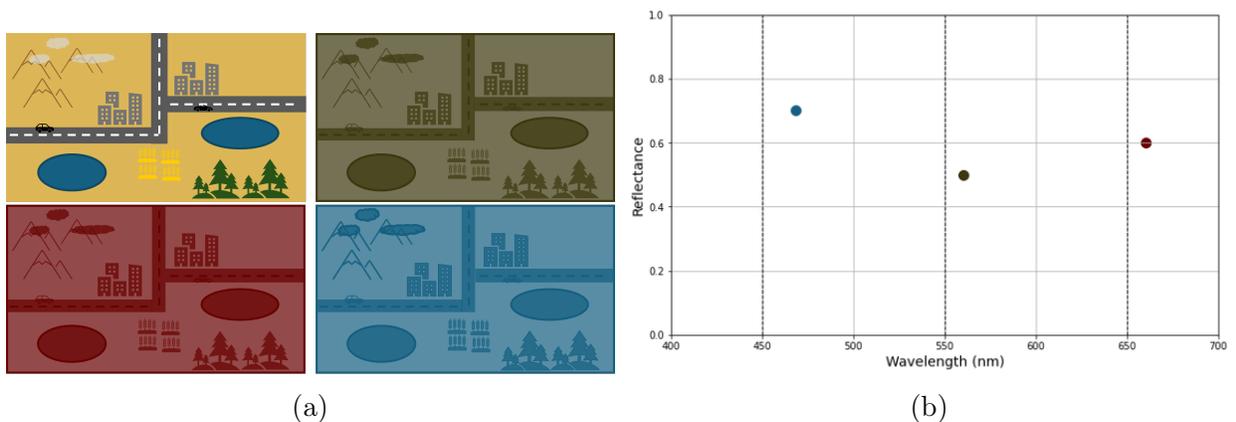


Figure 1.5: Display of the reflectance values of specific wavelengths corresponding to blue, green, and red color bands on a conceptual hyperspectral image.

While spectral sampling provides information about how materials reflect light across wavelengths, spatial sampling allows us to analyze the distribution of these materials in a given area. Specifically, spatial sampling refers to the selection of two-dimensional locations on the hyperspectral image, typically at the pixel level [31]. The ground sample distance (GSD) in an image is the distance on the ground between the centers of two neighboring pixels. GSD has also been used to represent the area that each pixel covers on the ground. Figure 1.2 illustrates spatial sampling on the conceptual image. The sensor embedded on the aircraft captures the regional data in the scene at the pixel level. If we imagine the region

on the ground corresponding to a pixel as a rectangle, the GSD corresponds to the length of the rectangle and is often thought of as a measure of pixel size or an inverse measure of spatial resolution. Each pixel represents a sample of light reflectance from a specific region in the image and can be used to identify substances in that location [31], as well as assist in the detection of similar materials in other areas of the image [32].

Figure 1.6 represents the results of spectral and spatial sampling. The left of the figure displays a three dimensional hypercube. Spatial sampling is accomplished by selecting a pixel or pixels on the x and y plane. Spectral sampling is indicated by the vector z and represents the spectral bands characterizing each pixel. The figure also shows a selected pixel capturing a soil region. This selected pixel has corresponding spectral information which is represented by the plotted spectrum.

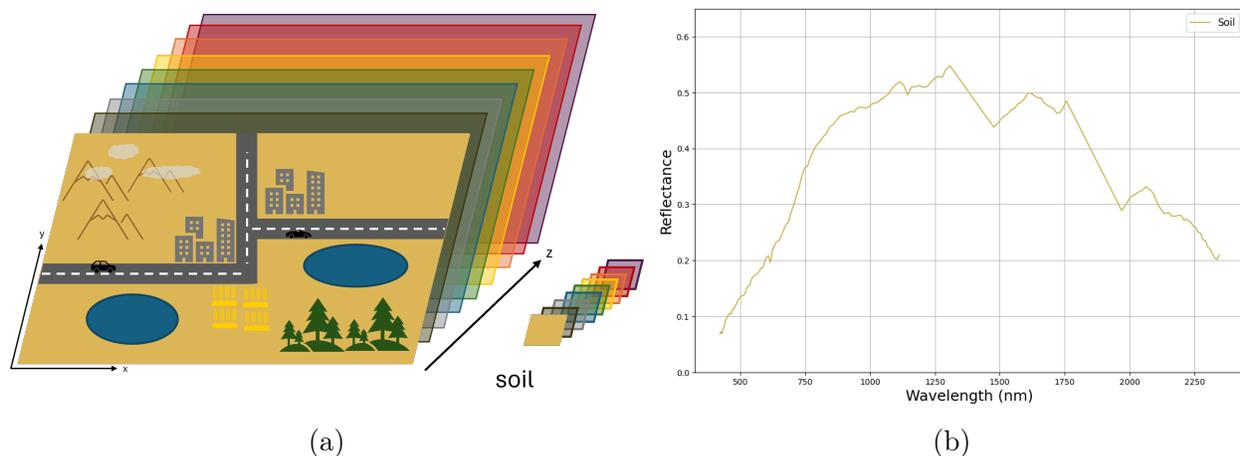


Figure 1.6: Depiction of the spectral signature from a soil pixel on a conceptual hyperspectral image. The plot illustrates the percent reflectance of light across a wavelength range from 400 to 2500 nm.

1.1.5 Spectral and Spatial Resolution Trade-off

Spectral resolution refers to the width — or number — of spectral bands used to measure material reflectance across the electromagnetic spectrum. Increasing spectral resolution involves narrowing the bands (across a set spectral range), which in turn provides finer detail and allows for more precision. Spatial resolution, on the other hand, defines the sensor's

ability to capture small distinct objects within a particular area [1]. Even though GSD can in certain scenarios be smaller than the sensor spatial resolution, the terms GSD and spatial resolution are often used interchangeably [31]. Increasing spatial resolution involves enhancing the sensor’s optics and reducing the pixel size, allowing researchers to capture finer details. While the aperture (the “eye” of the sensor) controls light intake, improvements in spatial resolution typically depend on enhanced sensor design and optical quality. Hyperspectral sensors typically have a higher GSD (larger pixel size) than multispectral sensors of similar construction and noise levels. Similarly, multispectral sensors have a higher GSD than comparable panchromatic (single band) imagers.

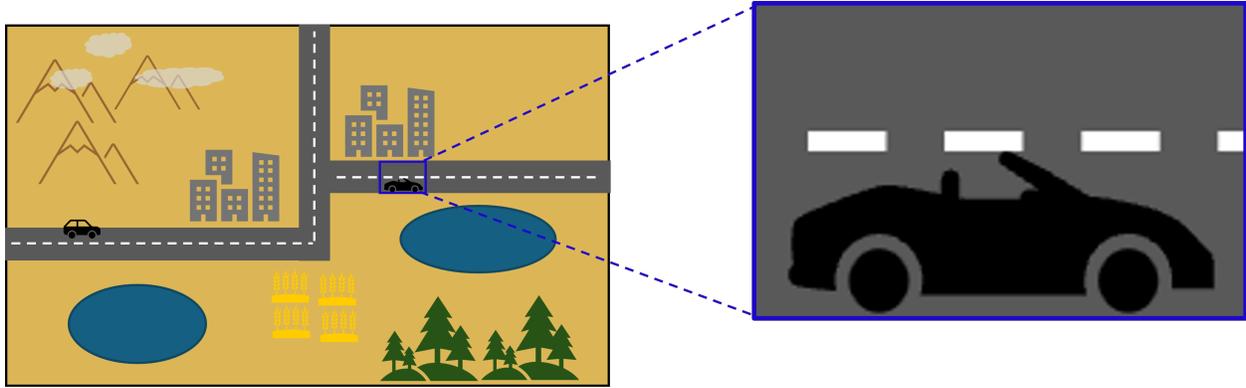
A unique trade-off between spectral and spatial resolution arises with the acquisition of hyperspectral images [4] — increasing the spectral resolution (adding more bands) vs. increasing the spatial resolution (more pixels per area). The signal-to-noise ratio (SNR) associated with hyperspectral data can decrease as the number of spectral bands increase, because the intensity of reflected electromagnetic radiation (the signal) is divided across more bands, resulting in reduced signal strength per band. Enlarging the pixel size or GSD can improve the SNR as larger pixels collect more light, leading to stronger signals. However, this increase in pixel size may result in lower spatial resolution of the image, making it harder to discern fine details. The primary use of hyperspectral imagery is to provide useful spectral information (high SNR, high spectral resolution). Thus, these images have a relatively low spatial resolution [4], [33]. Because the GSD of hyperspectral sensors is generally large — often due to the need for larger pixel sizes to accommodate the many narrow spectral bands resulting in lowered SNR — the spectrum measured from a pixel often represents a mixture of the spectra of the individual pure materials present.

1.2 Mixed Pixel Phenomenology

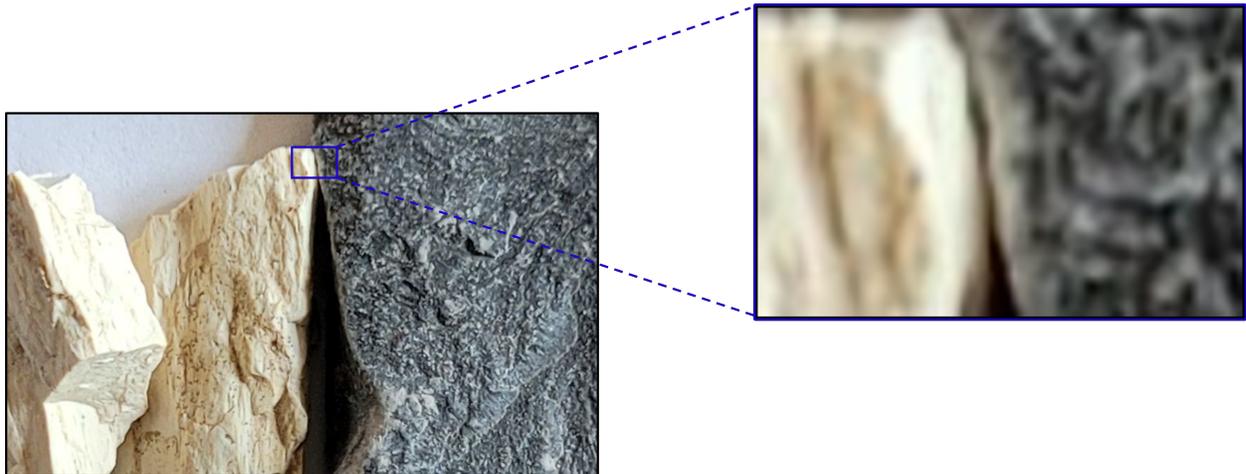
If only one material is present at the location measured by a pixel, the pixel is called a “pure pixel” and the spectrum for the pixel will be the spectrum for the given material [34]. However, the large pixel size associated with hyperspectral imagery as well as the phenomenology involved (i.e. atmospheric effects, material-to-sensor distancing, material-to-material proximity or a combination of these) imply that in most situations pure pixels are rare, and the measured spectrum is a mixture of the spectra of the individual materials [35].

The mixture of materials in the area represented by a pixel can be categorized as a linear or nonlinear mixture of spectra. Linear mixing occurs when the area measured by the pixel spectrum consists of one or multiple materials and light reflects off only one of the material at a time before reaching the sensor [36]. This is common when the materials occupy adjacent regions in the area covered by a pixel. An example of this scenario could be when the region on the ground measured in the pixel includes both a car and road surface around the car. Or as it pertains to this body of work, a pixel covers two distinctly separated minerals. Figures 1.7a and 1.7b display these examples, respectively.

Nonlinear mixing occurs when the measured pixel spectrum contains contributions from multiple materials, and the arrangement of these materials results in complex light interactions, such as multiple scattering or nonlinear reflectance effects [32], [33], [37]. There are various types of nonlinear mixture scenarios, but one that commonly occurs is when the individual materials are powders or grains, and the pixel area involves particles from multiple materials mixed together. In this arrangement, called a microscopic or intimate mixture [32], [33], light makes multiple bounces off of (and/or transmission through) different materials in the pixel area before reaching the sensor. Another nonlinear mixture can also occur when light passes through vegetation leaves and reflects off materials underneath before returning to the sensor, creating a canopy effect [33]. Figure 1.8 displays two mixture scenarios for



(a)



(b)

Figure 1.7: Presentation of two linear mixture examples: car on the road illustration (1.7a) and adjacent mineral composition (1.7b)

comparison. Figure 1.8a provides an example of a linear mixture of the minerals alunite and kaolinite. Conversely, an example of nonlinear mixing is shown in Figure 1.8b. To create the nonlinear example, the we ground the alunite and kaolinite samples shown in Figure 1.8a to obtain granular alunite (on the left in Figure 1.8a) and kaolinite (on the right in Figure 1.8a). Spectra of these granular materials were measured with a portable Analytical Spectral Device (ASD) field spectrometer — 2151 bands from 350-2500nm — and are shown in Figure 1.9 (left and center, respectively). We mixed the granular materials together vigorously. The resulting nonlinear-mixed spectrum shown is in the right-hand side of Figure 1.9.



Figure 1.8: Example of two types of mixture scenarios. Figure 1.8a is an example of a linear mixture of two materials. Figure 1.8b is a potential nonlinear mixture example where the components in the image are intimately mixed.

1.3 Data Sources

Hyperspectral data incorporates information primarily from two sources: spectral libraries and imagery from sensors and spectral libraries. The observed data is typically collected through advanced imaging systems, as discussed in section 1.1.1). One of the most widely used systems, and the one employed in this work, is the AVIRIS. Developed by NASA’s Jet Propulsion Laboratory (JPL), AVIRIS is designed to provide high-resolution spectral data across the VNIRSWIR ranges. The spectral features of the instrument currently capture data across 224 contiguous spectral bands, covering wavelengths from 380 to 2500 nm with each band approximately 10 nm wide, allowing for fine discrimination of material composition [38], [39]. AVIRIS balances the spatial and spectral resolution, delivering a high SNR that typically exceeds 1000:1 [40]. This means the actual signal (a surface’s reflected or emitted light) captured by AVIRIS is over 1000 times stronger than the background noise (e.g. sensor error, atmospheric effects, etc.)

The reference spectra, used to match the observed data, can be derived either directly

from the hyperspectral image or from pre-existing spectral libraries. Some of the most commonly used libraries include the ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) library and the United States Geological Survey (USGS) library [41], [42]. The ECOSTRESS spectral library — formally known as the Advanced Spaceborne Thermal Emission Reflection Radiometer (ASTER) spectral library — was created in 2018 in response to the launch of ECOSTRESS and its mission. The ECOSTRESS library contains all the spectra from the ASTER library, along with the addition of 1,116 vegetation spectra [41]. The ASTER library includes spectra of natural materials — e.g. snow, ice, rocks, minerals, lunar soils, terrestrial soils, and vegetation — and manmade materials from the JPL, Johns Hopkins University, and USGS [43].

Research Specifics

This body of work focuses on a subset of the AVIRIS image consisting of 400 x 350 pixels. Regions within the subset of the AVIRIS image were selected and employed as the observed pixel spectra. These observed pixels – regions of interest (ROIs) – were collected using the python package “Hyperspectralpy.” Using this package, we can export the selected pixels’ location in the image and associated spectral information for selected pixels to a .csv file [44]. The USGS library, specifically USGS Spectral Library version 7 (splib07a) [42] include the reference spectra employed in this research effort. The USGS splib07a spectral library contains 481 spectra of known mineral samples [38], [45].

1.4 Spectral Unmixing

Every observed pixel spectrum can be described by their pure material spectral signatures and associated fractional abundances [5], [46]. The abundance of each constituent spectrum in the pixel spectrum corresponds to the fractional abundance of area in the pixel occupied by the material [32], [47]. This process of identifying the pure materials in a pixel and

estimating material abundances is called spectral unmixing [33].

1.4.1 Linear Mixture Model

In line with different mixture compositions, spectral unmixing algorithms are generally characterized by two main approaches: linear mixture modeling (LMM) and nonlinear mixture modeling (NLMM) [32], [48]. Linear unmixing approaches assume the observed mixed pixel spectrum can be represented as a linear combination of pure materials and their fractional abundances [1], [33], [49]. More specifically, light photons are assumed to be reflected off only one material at a time before returning to the sensor [50]. Figures 1.7a and 1.7b depict cases where a LMM is appropriate. Ideally, the observed pixel spectrum from Figure 1.7a would be the linear mixture $\mathbf{y} = a_c \mathbf{s}_c + a_r \mathbf{s}_r$ where \mathbf{s}_c and \mathbf{s}_r are the spectra of the car and road, and a_c and a_r are the fraction of the area measured in the pixel occupied by the car and road, respectively [51]. It is important to note that the model representing the pixel shown in Figure 1.7b follows an identical structure. The standard LMM formula approximating the observed spectrum and representing a weighted sum of the spectra of individual materials [32], is shown below:

$$\mathbf{y} = \sum_{i=1}^N a_i \mathbf{s}_i + \epsilon = \mathbf{S}\mathbf{a} + \epsilon \quad (1.1)$$

where:

\mathbf{y} = observed spectral vector	\mathbf{S} = matrix of constituent spectra
\mathbf{s}_i = spectral signature of the i th material	\mathbf{a} = vector of fractional abundances
a_i = fractional abundance of the i th material	ϵ = additive noise vector (error term)
N = number of constituent spectra	

In Equation 1.1, \mathbf{y} is an $M \times 1$ vector representing an observed pixel spectrum, where M is the total number of spectral bands. Each \mathbf{s}_i is also a vector of size $M \times 1$, representing the

spectrum of an individual material present at the pixel’s location. The relative fractional abundances associated with each material is represented by the scalar value a_i . The terms, \mathbf{s}_i and a_i are summed across all N of the constituent spectra. In matrix notation, the linear mixture model becomes $\mathbf{S}\mathbf{a} + \epsilon$, where \mathbf{S} is an $M \times N$ matrix with each column corresponding to a spectral signature \mathbf{s}_i and each row corresponding to a specific spectral band. Each of the abundance values \mathbf{a}_i are components of \mathbf{a} , an $N \times 1$ abundance vector. The noise term ϵ is an $M \times 1$ vector representing the error values across the spectral bands.

The spectra \mathbf{s}_i are sometimes called endmembers, particularly in the context of linear unmixing when a single fixed set of \mathbf{s}_i — serving as the reference spectra — is derived directly from the image pixels. If the materials that are present in the pixel are known, then unmixing only requires finding the abundances that minimize the error in Equation 1.1. However, usually the set of materials in a pixel are not known prior to analysis (*a priori*), requiring the selection of \mathbf{s}_i from a potentially large library of reference materials. In this body of work, we will use the term endmember only for the spectra \mathbf{s}_i which are pixel spectra from an image. Since our emphasis is on unmixing with spectra of known materials from a library, we refer to the \mathbf{s}_i using the general term “constituent spectra.”

Delineation between endmembers can be somewhat subjective [32] and dependent on the analysis objectives. For example, if a pixel contains both soil and vegetation, there could be two end members in a scene. Alternatively, we could increase our enumeration to include the various minerals within the soil as well as the multiple types of vegetation also within the pixel [32]. Often by increasing the number of endmembers in a pixel, the unmixing problem can grow in complexity [33].

Extraction, Detection, Identification, Inversion, and Classification

Extraction, detection, identification, inversion, and classification are related terms regarding procedures associated with hyperspectral image processing and unmixing. Keshava et al. explain three primary steps involved in linear unmixing: dimension reduction, constituent

spectra determination, and inversion. These steps have been implemented simultaneously as well as consecutively in the unmixing process [33], [52]. Researchers often refer to the term extraction during the dimension reduction phase of unmixing. Due to the inherent, high dimensionality of hyperspectral data, feature extraction can be performed to remove unnecessary information and computational load of the preceding steps. However, this step is considered optional. Some examples of feature extraction steps include: variations of principle component analysis (PCA) and maximum noise fraction (MNF) [53].

An alternative use of the term extraction — regarding hyperspectral analysis — arises from endmember extraction. Many approaches (e.g. pixel purity index, N-FINDR, and iterative error analysis) have been developed to extract pure materials directly from the image pixel [5]. Detection and identification are associated with the constituent spectra determination phase. Researchers characterize detection as deciphering the spectra that are rare or uncommon in comparison to the surrounding spectral information. The identification step, in constituent spectra determination, is often executed after detection. In the identification stage, these anomalies are matched to one or many materials within a spectral library [51], [54]. Inversion is the final step in the unmixing process. The inversion phase refers to the process of estimating the abundance of the determined material in the scene. Image classification (in the context of this paper) is broadly defined as the ultimate goal of unmixing. Conversely, classification can also be referred to as the process of assigning material labels to pixels with the use of a classifier (e.g. a neural network) and training data [54].

1.4.2 Nonlinear Mixture Model

There are situations where the function used to approximate \mathbf{y} based on the constituent spectra \mathbf{s}_i (the right-hand side of Equation 1.1) is not linear. The process of determining the coefficients or parameters for this function is called nonlinear unmixing. Nonlinear unmixing approaches do not assume the observed pixel spectrum is generated by a linear combination of the endmembers. Instead, they account for more complicated interactions between light

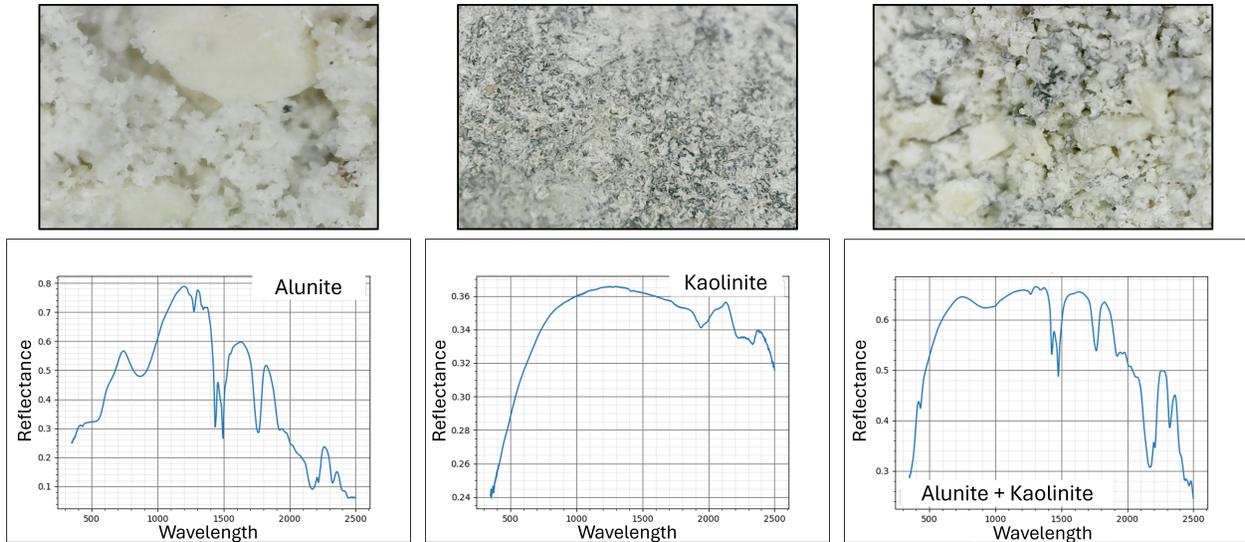


Figure 1.9: Display of granular mineral compositions: alunite, kaolinite, and a mixture of the two minerals (from left to right). Underneath each image is their associated pixel spectrum.

and materials [37]. The alunite-kaolinite intimate mixture shown in Figures 1.8b and 1.9 could be modeled using nonlinear unmixing techniques.

A polynomial function is the most common form of nonlinear unmixing, where higher order terms are used to model multi-bounce photon trajectories and other nonlinear interactions. The coefficients of these higher-order terms quantify the contribution of these interactions to the observed spectra. While the polynomial regression formula is nonlinear with respect to the constituent spectra \mathbf{s}_i , it is linear with respect to the individual terms, and thus readily solvable when the number of distinct constituent spectra is small. There are various other types of nonlinear mixture models such as bilinear models, the Hapke model and neural networks to name a few. Selection of approach depends on the composition of the pixel and the goal for unmixing. Though nonlinear models have the potential to achieve higher accuracy in abundance estimation, certain nonlinear models demand information that is not easily obtained by researchers [55].

Natural mixing can cause materials to become intimately mixed, making it challenging to distinguish endmembers and estimate abundances using linear methods [37]. Implementation

in certain scenarios (e.g. intimately mixed material) runs the risk of decreased accuracy in abundance estimation [56]–[58]. However, many of spectral unmixing studies focus on linear unmixing because of its faster computation and acceptable approximations for real world scenarios [32], [59], [60]. Manolakis et. al asserts that exact fractional abundance measurements are not necessary to assess model quality. “Goodness of fit” estimates — like mean squared error (MSE) or root mean squared error (RMSE) — suffice [51]. Thanks to research by authors like Combe et al., Heylen et al., Heinz et al., Schmidt et al., and Themelis et al., linear mixture modeling has been widely implemented in unmixing pixel compositions, even in cases where potential nonlinear interactions may be present [55], [56], [61]–[63]. The performance of the implemented linear techniques will also be compared to alternative NLMM approaches.

1.5 OLS for Unmixing

The LMM — depicted by Equation 1.1 — is often framed as a regression problem [33], [37] with OLS being a commonly implemented modeling method [64], [65]. In OLS regression, we estimate the model parameters by minimizing the sum of squared differences — also known as the residual sum of squares — between the observed and inferred reflectance values [64]. Alternative regression approaches exist (e.g. generalized or weighted least squares), but this body of work focuses on OLS because it serves as the baseline for most unmixing techniques [65]. In the OLS model, the observed mixed pixel spectrum is represented by the dependent variable \mathbf{y} , and can be further described as a vector of the pixel’s recorded reflectance across M spectral bands. Unmixing using this technique involves determining constituent spectra \mathbf{s}_i and their associated coefficients a_i that result in small error or very little difference between the observed and inferred spectra. The observed spectrum is modeled by inferring the constituent spectra from a spectral library, represented by the matrix \mathbf{S} . This spectral library is used to compute the unknown parameters, represented by the vector \mathbf{a} . When the

spectral library is small, we can solve for the abundances using matrix inversion. Equation 1.2 [52] displays this unmixing solution estimating the unknown, fractional proportion of materials.

$$\hat{\mathbf{a}} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{y} \quad (1.2)$$

All models inherently incorporate some level of error. Thus, the unmixing solution, shown in Equation 1.2, is an estimate because it excludes the error vector, ϵ . In least squares models, the error is referred to as the residual or in this context, the error is the distance between the inferred reflectance and the true pixel reflectance at wavelength j to M . Common metrics for assessing model error are MSE (presented by Equation 1.3) and RMSE (presented by Equation 1.4). MSE is calculated by normalizing the residual sum of squares by the number of data points (i.e., spectral bands). RMSE, calculated by taking the square root of MSE, is the error estimate utilized in this body of work.

The OLS objective is to minimize the error by identifying the set of coefficients that maximize the similarity between the response and the regressor values. Equation 1.5 defines this optimization problem by minimizing the RMSE between the observed spectrum and the inferred model. Assuming the matrix of constituent spectra \mathbf{S} is invertible, Equation 1.2 serves as the closed form solution to this optimization problem.

$$\text{MSE} = \frac{1}{M} \sum_{j=1}^M (\mathbf{y}_j - (\mathbf{S}\mathbf{a})_j)^2 \quad (1.3)$$

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{j=1}^M (\mathbf{y}_j - (\mathbf{S}\mathbf{a})_j)^2} \quad (1.4)$$

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \sqrt{\frac{1}{M} \sum_{j=1}^M (\mathbf{y}_j - (\mathbf{S}\mathbf{a})_j)^2} \quad (1.5)$$

1.5.1 OLS Assumptions

In spectral analysis, researchers often discuss the assumptions underlying the LMM, but the assumptions of OLS are rarely addressed. The outline below highlights primary OLS assumptions and discusses their validity in the context of spectral unmixing [65]–[67]:

1. Linearity in Relationships — OLS assumes the dependent variable is a linear combination of the predictor variables and the residuals. Specifically, linearity is exhibited through the additive nature between parameters. This relationship is modeled as a linear function (shown by Equation 1.1) in which both the contribution of the input variables and the error term are considered additive. The linear OLS objective function, defined in Equation 1.5, minimizes the residual sum of squares and implicitly assumes Euclidean distance between the observed and inferred spectra.

This linear assumption holds primarily for hyperspectral unmixing scenarios with macroscopic mixtures. Endmembers in these pixels occupy distinct or adjacent regions thus, we assume light photons interact with only one material before returning to the sensor.

However, this assumption fails when light photons interact with more than one surface before returning to the sensor. Scenarios where this assumption fails are when light interactions with material mixtures are more complicated (e.g. scenes with vegetation or mineral deposits) [33], [50].

2. Independent Identically Distributed (IID) Error — The residuals from OLS are assumed to satisfy three key properties: independence, normality with mean centered at zero, and homoscedasticity. Independence implies that there is no correlation between the residuals. The error of one observation (i.e. band 1 residual) will not be correlated with the error of another observation (i.e. band 2 residual). Residuals are assumed to follow a normal distribution with mean equaling zero. This means the residuals follow a

bell-shaped curve when plotted as a histogram, and their expected value is zero (shown by Equation 1.6). OLS is considered an unbiased estimator if the true value of the parameter is equal to its expected value which requires the expected value of the residuals to be zero. Homoscedasticity means the variance or spread of the residuals is constant across the observations, and because the residuals are also assumed uncorrelated, the corresponding covariance matrix can be represented by Equation 1.7. This covariance matrix, of size $M \times M$ (corresponding to the number of spectral bands), is diagonal, where the σ^2 represents the constant variance and I represents the identity matrix. This structure implies the errors are isotropic, meaning the errors are not computed in any one direction [68]. Additionally, because the covariance matrix is equal to the product of constant variance and the identity matrix, there is no correlation between the residuals and no correlation between the residuals and the input variables [51].

$$E(\epsilon) = 0 \tag{1.6}$$

$$E(\Gamma_\epsilon) = \sigma_\epsilon^2 I \tag{1.7}$$

In the context of the LMM, a multivariate normal distribution is assumed for the residuals with the mean centered at the observed pixel spectrum. Under this assumption, the likelihood function for the fractional abundances is given by Equation 1.8 [51]. This approximation is considered generally reasonable, but errors derived from the LMM do not always adhere to the properties of independence and homoscedasticity. Continuous and adjacent spectral bands often exhibit correlations, especially between bands in close proximity [69], [70]. Factors such as correlated bands, spectral variability, and spatio-temporal effects can disrupt homoscedasticity. As a result, residuals may exhibit fluctuating variances or, in some cases, follow entirely different distributions [3].

$$L(\mathbf{a}; \mathbf{y}, \mathbf{S}, \sigma^2) = \frac{1}{(2\pi)^{M/2} |\sigma^2 \mathbf{I}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{S}\mathbf{a})^T (\sigma^2 \mathbf{I})^{-1} (\mathbf{y} - \mathbf{S}\mathbf{a})\right) \quad (1.8)$$

3. Noncollinear Regressor Variables — OLS requires the regressor variables exhibit non-collinearity, meaning there is no linear dependency between the input vectors and no input vector is a linear combination of any of the other input vectors. When the input vectors are linearly independent, a unique solution can exist which is considered vital to obtain the closed form solution presented in Equation 1.2.

Noncollinearity is possible when there are more spectral bands than library spectra ($M \geq N$). However, when there are more library spectra than spectral bands a unique solution of abundances cannot be found. If $N > M$, inversion is impossible resulting in infinite solutions. This means that any number of abundance combinations can reproduce the observed spectrum (\mathbf{y}).

4. Correctly Specified Models — The input variables are considered fixed implying that the information available to describe the response variable to known *a priori*. We assume that the pertinent variables are included in the model or otherwise excluded. Additionally, the implementation of OLS regression assumes the inherent OLS assumptions are valid.

Unmixing requires either the observed image or a spectral library. Both of these approaches can challenge the correctly established modeling assumption as well as the preceding assumptions. When unmixing a pixel using the image, we must assume that pure pixels of individual materials are present and can be mapped to the mixed pixel, or that there exists a pixel with similar spectral properties, indicating identical materials and or mixtures. Alternatively, unmixing with a spectral library inherently assumes that the library contains materials that match those in the pixel we aim to unmix. Any one of the assumptions discussed can fail causing large estimated error or erroneous modeling conclusions.

1.5.2 Practical LMM Constraints

The unconstrained OLS model can lead to unrealistic abundance estimates, often resulting in negative values. This issue arises from overfitting the observed data and minuscule presence of certain substances in the observed spectra [71]. To address this, practical constraints are typically introduced. The two most common constraints applied are the nonnegative abundance constraint and the sum-to-one constraint [72]. The nonnegative abundance constraint, often referred to as nonnegative least squares (NNLS), ensures that all abundance values in the model are greater than or equal to zero. This constraint prevents the model from assigning negative values to the abundances of substances, which would be physically meaningless [55]. Equation 1.9 represents the nonnegative abundance constraint imposed on the OLS model, where N represents the number of explanatory variables or constituent spectra.

The other frequently applied constraint in LMMs is the sum-to-one constraint (shown in Equation 1.10). This constraint ensures that the estimated abundances of all substances in a pixel sum to one, aligning with the assumption that the total fraction of the materials in the pixel must equal 100% of the observed composition [33]. The combination of both the nonnegative abundance constraint and the sum-to-one constraint is often referred to as the fully constrained model. While the sum-to-one constraint is relatively simple to implement and reflects the physical requirement of a complete pixel composition, it may not always be appropriate. In some cases, it can be relaxed, especially when “complete prior knowledge” of pixel components is unavailable [52], [72], [73]. In this body of work, we relax the sum-to-one constraint based on the assumption that some materials may appear brighter under some conditions, causing the estimated abundances to exceed one. Therefore, we only enforce the nonnegative abundance constraint on our models.

$$a_i \geq 0 \quad \forall 1 \leq i \leq N \tag{1.9}$$

$$\sum_{i=1}^N a_i = 1 \tag{1.10}$$

1.5.3 Challenge with OLS Unmixing

Pure materials can exhibit different spectral characteristics due to various factors. These factors include fluctuations in lighting conditions, particle size, surface roughness, and atmospheric influences [4], [55]. This means that both different materials and the same material can display different spectral patterns. For example, minerals from some chemical categories will have inconsistencies in appearance due to their proximity to minerals in other chemical categories. Ultimately, the challenge of spectral variability creates a need for extensive spectral libraries to improve material identification [74]. Solving for the coefficients in a least squares regression problem with a large spectral library — when there are more spectra in the library than bands — results in an ill-posed inference problem [74] requiring the inversion of a non-invertible matrix.

As spectral libraries have become more common, unmixing with spectra from a library rather than image pixels has become substantially more important in practical applications. Unmixing with library spectra provides precise information that may be essential for the user (e.g. the specific mix of minerals present for a geological understanding of an area, or a unique chemical pollutant that may be present in sand or soil.) As such, unmixing with a library provides a process for sub-pixel material identification [75].

When dealing with microscopic mixture scenarios, the endmembers are typically not known *a priori*, therefore an extensive library of materials is often used for unmixing. OLS regression often fails to accurately detect constituent spectra when spectral libraries are large. Figure 1.10 depicts an attempt to unmix a pixel from an image over the Cuprite Hills mining site using OLS. The inferred pixel spectrum includes spectra from the entire USGS mineral spectral library — represented by the non-bold colored spectra — to model the observed pixel spectrum. Modeling accuracy is presented through the fit of the plotted observed and

inferred spectra and assessed using RMSE. In the plot, the observed and inferred spectra have a very close fit. However, this result is erroneous and impractical because the entire spectral library of materials is not present in the pixel nor can negative abundance values exist (which was needed to achieve this close model fit) for materials present in the pixel.

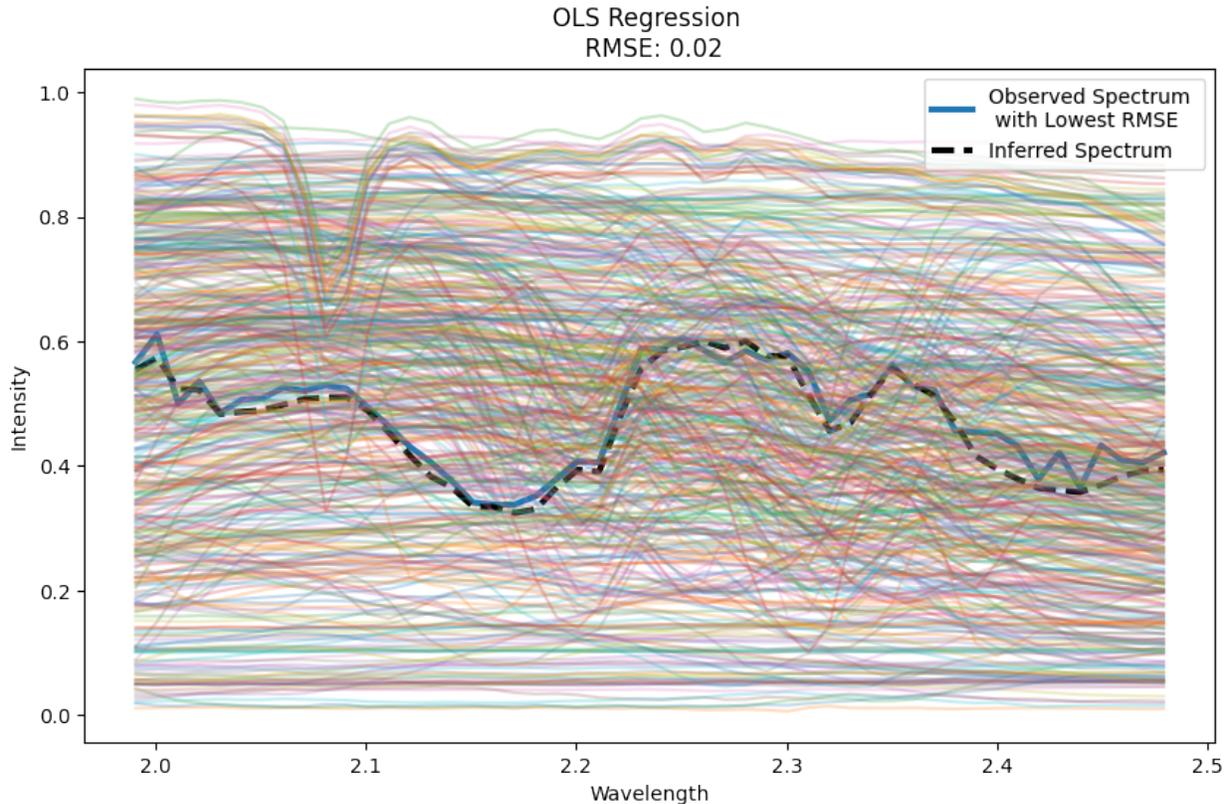


Figure 1.10: Display of OLS unmixing results. The bold blue spectrum represents the observed pixel. The dashed spectrum represents the inferred model. The non-bold colored spectra represent the material included in the inferred model.

The increase in library size and complexity is not simply an academic problem. For example, soil has historically been considered a pure endmember category in the image unmixing paradigm, but soil is in fact very complex [76], [77]. As of 2016, the “global vis-NIR soil spectral library” [78] contained over 23,632 distinct spectra measured across 350 - 2500 nm from 92 different countries, and almost all spectra contain substantial metadata. The scale, variety, and specificity of libraries like this are outpacing exploitation methods and the collective understanding of how to utilize these methods on such large and detailed

VNIRSWIR spectroscopy data.

1.6 Alternative Unmixing Techniques

1.6.1 Regularization and Sparse Approaches

Although the linear mixing model is relatively simple, this model has fostered a huge amount of research leading to “a plethora of unmixing algorithms” [32]. LMMs used for unmixing intimate or nonlinear mixtures often benefit from constraints, regularization, or Bayesian techniques. Regularization and sparse approaches are often needed in spectral unmixing because they address some of the assumption failures associated with OLS regression. Regularization is a technique used to prevent overfitting by enforcing a penalty on the complexity of the model. Overfitting prevention ensures that the solutions proposed are robust and generalizable. High correlation among library spectra and between bands introduces a notable challenge in spectral unmixing. Large spectral libraries are convenient for capturing spectral variability, but their use can lead to overfitting, making regularization necessary. Moreover, extensive libraries can complicate the inversion process. Some linear unmixing techniques inherently incorporate regularization through penalties on model complexity, eliminating the need for dimensionality reduction methods.

The common penalties imposed on the model coefficients are the L_1 norm (also known as Least Absolute Shrinkage and Selection Operator, or LASSO regression), L_2 norm (commonly referred to as ridge regression) and a combination of the L_1 and L_2 norm (ElasticNet) [79]. LASSO and ridge regression are extensions of OLS with the L_1 and L_2 penalty norms weighting the material abundances, respectively [80], [81]. The L_1 norm incorporates Manhattan distancing while the L_2 norm incorporates euclidean distancing [82]. Additionally, it is important to recognize that these regularization techniques are related to Bayesian regression: LASSO regression is Bayes optimal when the parameters follow a Laplace prior distribution, whereas ridge regression is Bayes optimal when the parameters adhere to a

Gaussian prior distribution [80].

The L_2 norm penalty applied to the abundances results in shrinkage of the values [82], making them smaller but not necessarily zero. As a result, even within a positively bounded solution space, the fractional abundances can be very small percentages but will never reach zero. This regularization characteristic leads to large model sizes, with many constituent spectra included in the estimated spectrum, ultimately causing substantial overfitting of the observed spectra. Equation 1.11 represents the objective function of ridge regression, where $\hat{\mathbf{a}}$ denotes the estimated abundance vector. Note the double bars ($\|\cdot\|_2^2$) represents the squared L_2 norm replacing the summation symbols, from the previous equations. The goal of ridge regression is to minimize the expression on the right-hand side, consisting of two components: the residual sum of squares and the model coefficients. The first term, $\|\mathbf{y} - \mathbf{S}\mathbf{a}\|_2^2$, represents the residual sum of squares, which quantifies the difference between the observed spectrum \mathbf{y} and the inferred model $\mathbf{S}\mathbf{a}$. The second term, $\lambda\|\mathbf{a}\|_2^2$, introduces the L_2 penalty. This penalty shrinks the size of the abundance estimates to prevent overfitting. Specifically, λ is a regularization parameter that controls the strength of this penalty. A higher value of λ results in a stronger penalty on the abundance estimates.

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \left(\frac{1}{M} \|\mathbf{y} - \mathbf{S}\mathbf{a}\|_2^2 + \lambda \|\mathbf{a}\|_2^2 \right) \quad (1.11)$$

While ridge regression is a form of regularization, it may still retain many variables that do not significantly contribute to the model. Thus, sparse unmixing emerges as a valuable alternative by selecting an optimal subset of input variables from the library to more accurately describe the observed spectrum [4]. Numerous studies have achieved success using sparse unmixing approaches. Typically, sparsity is enforced through the use of a penalty which shrinks insignificant coefficients or the abundances of absent substances. LASSO regression shrinks coefficients to zero, thereby penalizing the size of the model [80], [82]. Equation 1.12 displays the modified version of Equation 1.11 which incorporates the L_1 penalty on the model coefficients. The LASSO regression objective function promotes sparsity by forc-

ing many coefficients to exactly zero, effectively excluding variables that do not contribute meaningfully to the model. Subsequently, the L_1 penalty adheres to the nonnegative abundance constraint as the coefficients are driven toward zero but cannot fall below it. The set of coefficients $\hat{\mathbf{a}}$ minimizes the objective function subject to the L_1 penalty norm.

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \left(\frac{1}{M} \|\mathbf{y} - \mathbf{S}\mathbf{a}\|_2^2 + \lambda \|\mathbf{a}\|_1 \right) \quad (1.12)$$

ElasticNet is a technique that balances the L_1 and L_2 norms, combining benefits of both regularization types. The ElasticNet objective function is shown by Equation 1.13. In this Equation, λ_1 and λ_2 are the regularization parameters, often selected via cross validation and set between 0 and 1, that control the influence of the L_1 and L_2 norms. The λ_1 governs the sparsity of the solution, while λ_2 manages the shrinkage of less influential coefficients.

The combination of the L_1 and L_2 norms has been applied in spectral unmixing across several studies. Though not Elasticnet, Li et al. proposed a new algorithm incorporating the $L_{2,1}$ norm to preserve spectral information for pixels in the same region [83]. Iordache et al. presented a comparative study of an algorithm using the $L_{2,1}$ norm, CLSUnSAL — originally introduced by Bioucas-Dias and Figueiredo — alongside other established sparse regression techniques [81], [84]. We implemented ElasticNet in an effort to select the subset of primary spectra (model features) from the library [85]. Additionally, the flexibility of ElasticNet allows for easy adaptation to different modeling needs by adjusting the relative influence of each regularization type.

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \left(\frac{1}{M} \|\mathbf{y} - \mathbf{S}\mathbf{a}\|_2^2 + \lambda_1 \|\mathbf{a}\|_1 + \lambda_2 \|\mathbf{a}\|_2^2 \right) \quad (1.13)$$

1.6.2 Iterative Approaches

An alternative to sparse approaches is the use of iterative techniques, such as optimization strategies like breadth-first and depth-first search strategies. In the traditional sense,

breadth-first search (BFS) explores multiple variable paths concurrently, evaluating all potential variables at each step. In contrast, depth-first search (DFS) delves deeply into one path before searching another path, exhaustively exploring one variable or path before returning to explore others. In the context of spectral unmixing, BFS and DFS approaches resemble iterative processes that prioritize variables based on their significance. BFS evaluates a broad set of potential spectra at each step, aiming to find a balance between exploration and model accuracy. DFS examines the significance of one spectrum at a time which in turn can be useful in identifying a smaller subset of influential variables. Both strategies are applied iteratively to optimize model selection, making them suitable for identifying primary variables in spectral unmixing.

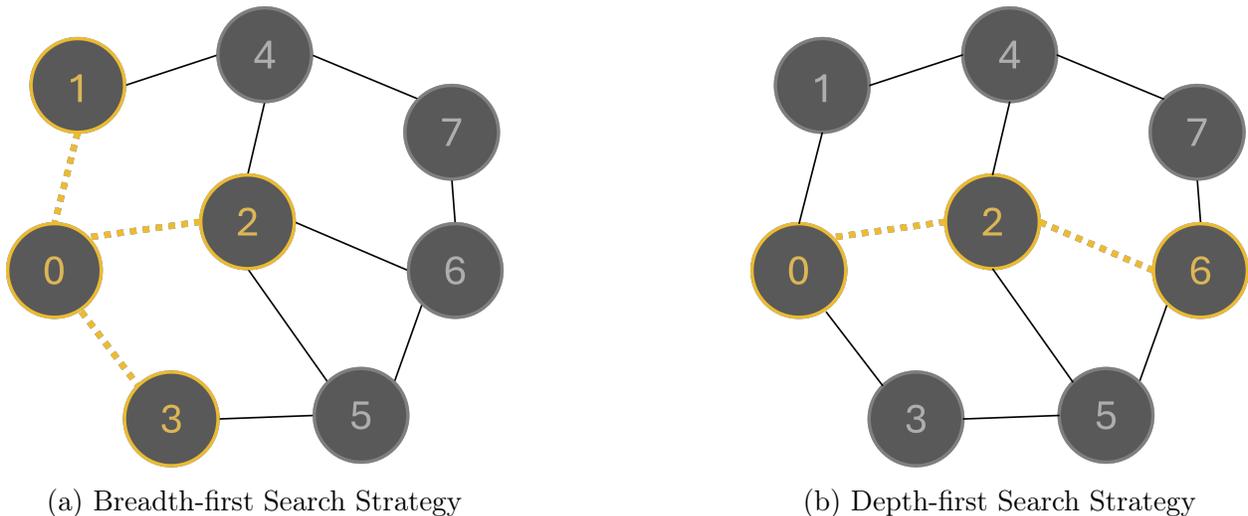


Figure 1.11: Comparison of search optimization strategies: Breadth-first search (a) and Depth-first search (b).

A closely related technique to this form of model selection optimization is stepwise regression. However, certain nuances of stepwise regression may not fully align with spectral unmixing assumptions, particularly regarding multicollinearity and negative coefficients. It is for this reason, we refer to the iterative approaches incorporated in this body of work as breadth- and depth-first search optimization.

Stepwise regression is primarily implemented for model selection with a secondary aim of parameter estimation. In this context, stepwise regression directly responds to the goal

of sparse unmixing. Specifically, we are identifying the subset of explanatory variables — from a far larger set of materials — that provide the best fit for the observed data [86], [87]. However, concerns over multicollinearity limit the use of stepwise regression’s use in some cases. When explanatory variables are highly correlated, stepwise regression models struggle with deciphering variable influence on the response variable. R.D. Routledge says that in certain scenarios of high correlation, the combination of explanatory variables can show a significant impact on the response variable, while individually the variables are insignificant [88]. Additionally, in another paper by Michael Lewis-Beck, the author explains an alternative scenario where three variables are considered significant individually, but as they are iteratively included in the model, their significance lowers. He claims the iterative process fails to depict the relationship between the variables and ultimately their level of significance [89].

The limitations of stepwise regression often exhibit two characteristics that spectral unmixing does not. Many of the correlated variables in the spectral library represent variations of the same substance rather than two distinct, highly correlated variables. Additionally, OLS variables can negatively impact the response variable, which complicates the interpretation of abundance values because negative coefficients are impractical. Despite these limitations, the search strategies employed by stepwise regression — particularly its iterative nature — offer advantages when adapted to spectral unmixing.

Considering the attributes of spectral unmixing, breadth-first and depth-first search strategies provide a more effective solution. Breadth-first and depth-first search strategies iteratively identify the primary variables, and aim to compute the most statistically significant substances influencing the observed pixel reflectance. Winter et al. explains that the iterative process of stepwise regression coupled with nonnegative abundance enforcement lowers modeling error [71]. In another paper, Gault et al. compares NNLS, sparse, and stepwise approaches. The authors explain that stepwise regression can out perform non-iterative approaches especially when detection is the main goal [90].

Breadth- and depth-first search optimization share similarities with stepwise regression because these approaches iteratively identify the most significant explanatory variables and incorporate them into the model. The significance of each variable is evaluated using the F-statistic, a ratio that compares explained versus unexplained variance. The F-statistic is used to calculate the probability (p-value) of obtaining the observed ratio by chance. A low p-value suggests that the computed F-statistic is unlikely to have occurred randomly. When the p-value falls below a predetermined threshold, the variable is considered to contribute significantly to the variance observed in the response variable and subsequently included in the model [87], [88]. Ultimately, breadth-first and depth-first optimizations align well with unmixing goals, as they systematically address variable selection to produce the best-fit model, optimizing both accuracy and interpretability in spectral data contexts. We, subsequently implement breadth-first and depth-first optimizations as the iterative model selection approaches within this body of work.

1.6.3 Mixed Integer Nonlinear Program (MINLP)

Mathematical programming is another powerful technique designed to identify optimal solutions by adjusting decision variables (e.g., the spectra included in the model) within established boundaries (e.g., nonnegative abundances or sum-to-one conditions) that represent practical limits such as physical properties or resource constraints. Despite its potential for exact abundance solutions [91], few researchers employ mathematical programming for spectral unmixing. Among the studies that have utilized mathematical programming, there is a noticeable lack of consistent terminology. Many papers describe techniques that originate from the broader field of mathematical programming, such as linear programming [92], [93], integer programming [94], mixed integer programming (MIP) [95], [96], MINLP [94], and other variations [97]–[99]. While these techniques are related, each employs slightly different approaches. Mathematical programming is a field of study that analyzes optimal decision making subject to a set of limitations. Solutions are derived by optimizing an objec-

tive — a minimization or maximization function — composed of decision variables. These variables are iteratively adjusted to find the best solution within the defined constraints. MINLP, in particular, is an optimization method that extends MIP by incorporating nonlinear relationships. This includes combining concepts from both linear, integer and nonlinear programming.

1.6.4 Bayesian Model Averaging (BMA)

BMA is a technique in which we ensemble a multitude of models. In Bayesian statistics, probability is used to quantify uncertainty about parameters and models, updating these beliefs as new data is observed. The posterior probability is an important component in Bayesian inference that represents the updated probability of a model after observing data. BMA was initially developed as a method to address model uncertainty in statistical inference. The idea of averaging across models using bayesian methods emerged in the 20th century. However, BMA was formally established by David Madigan and Adrian Raftery. The authors assert that selecting a single “best” model from a set often ignores the uncertainty inherent in the model selection process. They proposed averaging across models using weights based on each model’s posterior probability [100].

In the context of spectral unmixing, BMA involves the combination of hundreds of mixture models with the primary aim of estimating abundances. Equation 1.14 shows a basic BMA formula in which N is the total of number models, $\hat{a}^{(n)}$ represents the estimated vector of abundances given by model M_n , and $P(M_n|\mathbf{y})$ depicts the posterior probability of model M_n given the observed pixel spectrum.

$$\hat{\mathbf{a}} = \sum_{n=1}^M P(M_n|\mathbf{y})\hat{a}^{(n)} \quad (1.14)$$

1.7 Overview to Dissertation

1.7.1 Research Gaps

In the field of Hyperspectral Image Analysis, many papers develop new and useful approaches for material identification and classification, but far less research has been completed that evaluate technique performance in relation to comparative models. OLS has many extensions, but few unmixing papers provide a holistic performance comparison of its many adaptations. The few papers comparing OLS extensions consider a small subset and lack a comprehensive view of algorithm trade-offs, including responsiveness to the aforementioned challenges. Additionally, at this current time, there are no studies that provide this comparison in terms of physical-chemical feature detection of material types.

1.7.2 Contribution Outline

This research focuses on supervised, hyperspectral mixture modeling with the purpose of material identification irrespective of the substance type or pixel composition complexity. We assume successful unmixing for microscopic or intimate mixtures will also prove successful for macroscopic mixtures. Thus, we concentrate our efforts on unmixing mineral compositions. The techniques included in this body of work are shown in Table 1.1. This study has the concurrent aim of responding to the aforementioned challenges and are subsequently addressed through the incorporated techniques. Additionally, certain aspects of spectral variation will be explored with the purpose of determining the factors associated with specific materials that contribute to successful and unsuccessful material identification.

Within the scope of this dissertation, some key nuances to this research are:

1. In Chapter 1, we explain the assumptions behind OLS and its relevance to spectral unmixing.

2. In Chapter 2, we provide a thorough comparison of unmixing with unconstrained least squares and its extensions or alternative interpretations [101].
3. In Chapter 2, we develop an algorithm taxonomy demonstrating the relationships and nuances among approaches [101].
4. In Chapters 2, 3, and 4, we provide a framework for benchmarking unmixing techniques i.e. feature recommendation metrics [101]–[103]. Appendix A showcases — in a full table — the tradeoffs of each of the incorporated techniques for each ROI within this body of work.
5. In Chapter 4, we develop a novel approach addressing misalignments of OLS assumptions with spectral unmixing [103].
6. In Chapters 3 and 4, we provide a discussion and taxonomy of the physical-chemical material phenomenon advancing successful unmixing [102], [103].

Unmixing Techniques Categorized by Approach

Non-sparse	Ordinary Least Squares Regression Nonnegative Least Squares Regression Ridge Regression
Sparse	LASSO Regression ElasticNet
Optimization	Depth-first Search Feature Selection - NNLS Depth-first Search Feature Selection - LASSO Depth-first Search Feature Selection - ElasticNet Breadth-first Search Feature Selection - NNLS Mixed Integer Nonlinear Programming
Ensambling	Bayesian Model Averaging - NNLS Bayesian Model Averaging - LASSO Bayesian Model Averaging - ElasticNet Quadratic Bayesian Model Averaging - NNLS Quadratic Bayesian Model Averaging - LASSO Quadratic Bayesian Model Averaging - ElasticNet
Band Decorrelation	HySUDeB

Table 1.1: Categorization of implemented techniques by unmixing approach

2 Theoretical and Practical Progress in Spectral Unmixing from a Sparse Perspective

2.1 Introduction

In this chapter, we provide a thorough performance evaluation of the following investigative methods: OLS regression, NNLS, ridge regression, LASSO regression, feature search strategies and BMA. These unmixing approaches were evaluated using multiple criteria: incorporation of nonnegative abundances, model size, accurate mineral detection, and RMSE. We provide a taxonomy of the regression methods, showing that many methods can be understood as Bayesian methods with specific priors.

2.2 Methods

2.2.1 Data Collection

We collected the observed pixels — regions of interest (ROIs) — using the python package “Hyperspectralpy”. This package exports the selected pixel locations (in the image) and associated spectral information to a .csv file [44]. The alunite ROI was comprised of 123 pixels and the kaolinite ROI was comprised of 120 pixels. Figure 2.1 shows the ROIs within the dataset: alunite and kaolinite are indicated by the orange and purple shaded regions, respectively. The two ROIs were selected with the goal of successfully detecting the primary

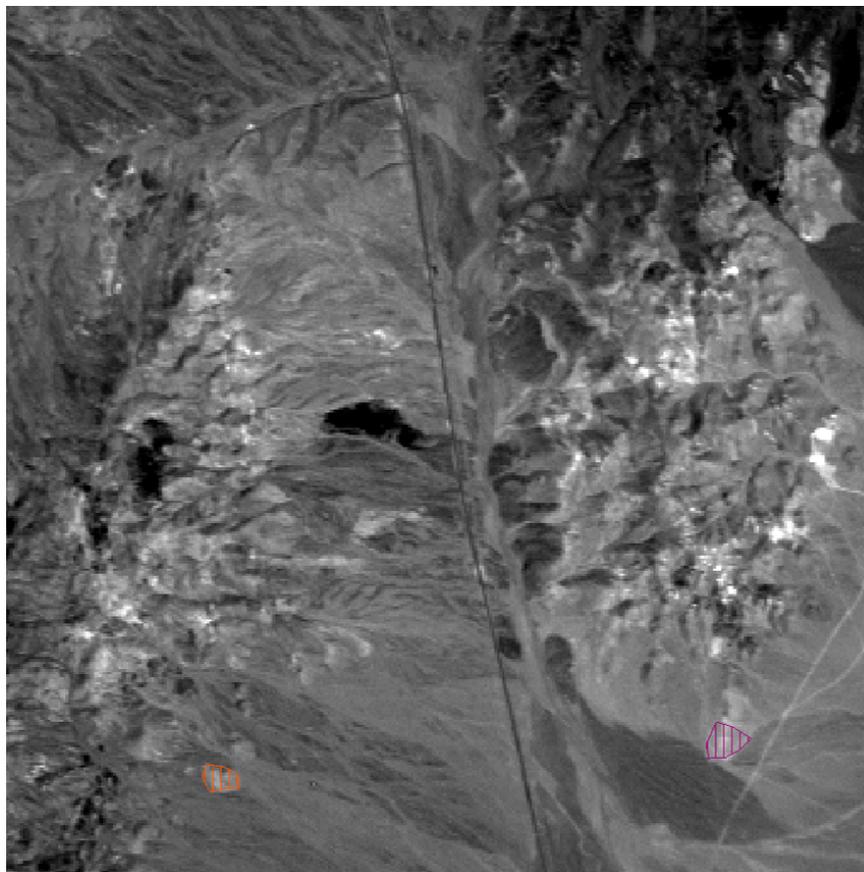


Figure 2.1: Display of the region of selected pixels using “Hyperspectralpy” [44]. The shaded orange region indicates the selected alunite pixels and the shaded purple region indicates the kaolinite pixels.

mineral associated with that pixel location (alunite and kaolinite). Table 2.1 displays the unmixing techniques incorporated in this study as well as their RMSE functions.

2.2.2 Unmixing Algorithms

We incorporated the OLS unmixing results in this study to present a holistic glimpse surrounding the progression of this research. The unconstrained OLS model means the model had no requirements associated with constraining the material abundances. Material abundances in this model could assume negative values and a sum greater than one. We utilized the sklearn “LinearRegression” package to implement the unconstrained OLS unmixing [104]. In this study, the NNLS technique solely constrains the abundances to be nonnegative. We

avoided constraining the abundances to sum less than or equal to one. To incorporate the NNLS technique, we implemented the “NNLS” python package from `scipy.optimize` [105]. We implemented ridge regression and LASSO regression individually to uncover any unmixing strengths related to the individual approaches. “Ridge” and “Lasso” python packages from `sklearn` were used for implementation [106], [107]. The alpha parameter was set to 0.0005 for LASSO regression and the default value (alpha = 1.0) for Ridge regression. We implemented two iterative approaches — DFS and BFS strategies — to ascertain the unmixing advantages each model provides. DFS and BFS are commonly described in terms of graph or tree based analogies, but these strategy can also be effective in unmixing. With DFS, we start with an empty set and incrementally increase the model size based on predefined criteria — specifically, a p-value threshold of 0.05 in this study — for updating the model. Alternatively with BFS, the entire spectral library is included in the set of model features and we decrease the model size based on the predefined criteria. The code for DFS and BFS were developed independently rather than utilizing a python package. Both techniques are an extension of OLS regression in which we iteratively update the minerals incorporated as features in the model with the goal of minimizing error. We conducted NNLS regression on the model set of minerals deemed most significant following the DFS and BFS searches to estimate the fractional abundances. In this study, two techniques for BMA were explored — a basic model and quadratic (BMA-Q) model. We experimented with second order terms to determine if higher accuracy can be achieved through modeling the interaction between two materials, in our implementation of BMA-Q. We coded both BMA and BMA-Q techniques independently rather than using a python package.

2.3 Results

Table 2.2 and Figure 2.2 present the comprehensive results and comparison for Alunite Hill 1 and Kaolinite Region 1. In Table 2.2, the performance of each technique is summarized in

Unmixing Technique	RMSE Function
Ordinary Least Squares (OLS)	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2}$
Nonnegative Least Squares (NNLS)	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2}, a \geq 0$
LASSO	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2} + \lambda \sum_{i=1}^N a_i $
Ridge Regression	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2} + \lambda \sum_{i=1}^N a_i^2$
Breadth-first Search	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2}, a \geq 0$
Depth-first Search	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2}, a \geq 0$
Bayesian Model Averaging	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2}, a \geq 0$
Quadratic Bayesian Model Averaging	$\sqrt{\frac{1}{M} \sum_{j=1}^M (y_j - (Sa)_j)^2}, a \geq 0$

Table 2.1: Overview of the unmixing techniques with their associated RMSE loss functions.

Alunite (1)	OLS	NNLS	RIDGE	LASSO	DFS	BFS	BMA	BMA-Q
Model Size	481	8	133	4	2	3	3	3
Detection Rate	1.0000	0.9675	0.9593	0.9593	0.9268	0.3252	0.8862	0.8455
Mean RMSE	0.0112	0.0250	0.2200	0.1348	0.0363	0.0672	0.0519	0.2743
Run Time (sec.)	0.0084	0.0027	0.0098	0.0070	1.8757	4.1270	0.0858	0.1355
Kaolinite (1)	OLS	NNLS	RIDGE	LASSO	DFS	BFS	BMA	BMA-Q
Model Size	481	8	187	6	2	4	4	5
Detection Rate	1.0000	1.0000	1.0000	0.8750	0.7833	0.4417	0.6667	0.6167
Mean RMSE	0.0102	0.0284	0.1251	0.1472	0.0412	0.0482	0.0630	0.3863
Run Time (sec.)	0.0088	0.0028	0.0101	0.0146	2.3485	4.0965	0.1255	0.2862

Table 2.2: Displays the comprehensive results of the unmixing techniques for the alunite and kaolinite ROIs.

terms of model size, target mineral detection rate, error and computation time. Each of these metrics are a calculated average across all the pixels in the ROI. Also, the target mineral detection rate refers to the number of inferred models that include alunite or kaolinite from the observed pixels in their respective ROI. Figure 2.2 displays the average model error on the y-axis, average model run time on the x-axis and the average model size by the diameter of the plotted dots. Specifically, the plotted techniques with small dot diameter and in the

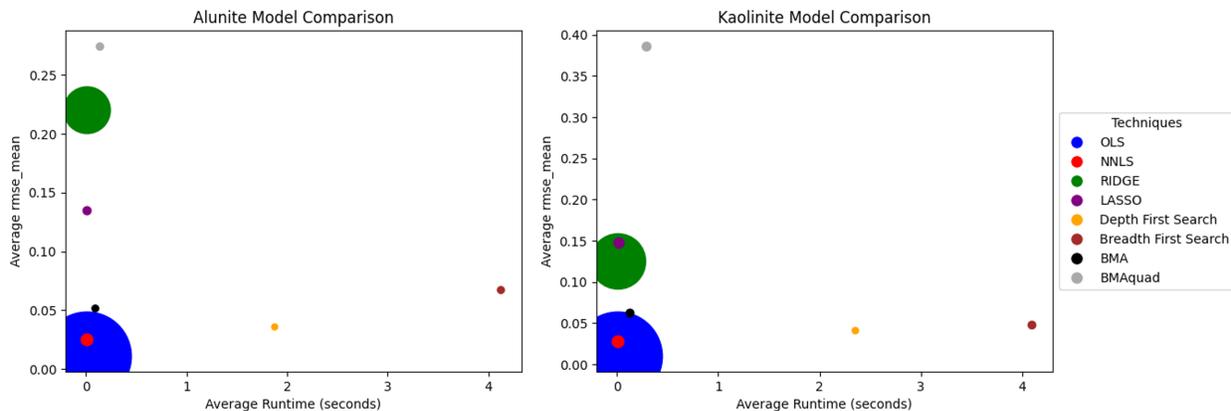


Figure 2.2: Comparison of the unmixing results based on RMSE, runtime, and model size. The left plot displays the results for alunite. The right plot displays the results for kaolinite.

bottom right corners of the graphs in Figure 2.2, achieved the highest performance.

OLS, ridge regression and BFS demonstrated the poorest performance in unmixing the range of pixels for both Alunite Hill 1 and Kaolinite Region 1. Though OLS had a relatively fast run time and low average error (RMSE = 0.01 for both ROIs), this technique incorporated negative abundance values in the models and included all spectra from the library in its models. Ridge regression performed similarly to OLS with large model sizes (showing evidence of overfitting within the inferred y-values). Ridge regression also introduced much higher average error, with mean RMSE values of 0.2 in the alunite ROI and 0.1 in the kaolinite ROI. BFS produced smaller model sizes, averaging three to four minerals included in the inferred models, and achieved lower error compared to ridge regression. However, its target mineral detection rate was below 50% for both ROIs. Additionally, BFS had the highest computation time among all techniques, averaging four seconds per model.

The BMA approaches achieved lower target detection rates and higher computation time than NNLS, and LASSO. BMA and BMA-Q differed in terms of error performance: BMA demonstrated relatively low error with RMSE values of 0.05 for the alunite ROI and 0.06 for the kaolinite ROI, whereas BMA-Q recorded the highest average RMSE across all techniques.

DFS computed small model sizes, enforced the nonnegative abundance constraint and maintained low average RMSE. Unlike BFS, DFS had higher detection rates of the target

minerals. DFS detected alunite in 93% of the pixels and kaolinite in 78% of the pixels. The computation time for each pixel was lower than BFS but still higher than the rest of the techniques.

Despite its sparse characterization, Lasso regression exhibited higher average model sizes than many of the other techniques — four minerals for the alunite ROI and six minerals for the kaolinite ROI — as well as greater average error compared to DFS and BFS. However, it achieved higher target mineral detection rates (above 80%) while maintaining substantially faster computation times.

NNLS achieved the second lowest average RMSE and the lowest run times. Also, using this technique, we were able to detect the target minerals in over 95% of the pixels for both ROIs. The average model size —eight minerals — was smaller than those of OLS and ridge regression but larger than all other techniques.

2.4 Discussion

This study aimed to unmix intimate mixtures using a variety of OLS-based techniques. OLS is a common interpretation of the linear mixture model but can result in impractical solutions or a non-invertible matrix from large spectral libraries. A common challenge associated with intimate mixtures is the high within class and external class variability from the image and thus a need for a large spectral library. We presented methodologies that modify the OLS technique to ensure positive abundance values, penalize model size and or ensemble multiple methods.

Besides OLS, Ridge regression and BFS performed poorly. The L2 penalty norm associated with ridge regression is designed to shrink the abundance values to become smaller, but not necessarily zero. This regularization attribute caused model sizes to remain large and substantially overfit the observed spectra. BFS had low detection of the target mineral and high computation time (due to its iterative methodology). Unmixing one pixel provides

fairly quick results but because we calculate the p-values for hundreds of models (for each pixel in the ROI), the computation time compounded. The low detection rate observed with BFS arises from its process of eliminating variables based on their perceived significance. Starting with the entire library incorporated in the model (essentially an OLS configuration), the contribution signals of individual minerals often become conflated. As a result, models including the target, may be removed because their significance is overshadowed or goes unrecognized.

We achieved the best unmixing results with NNLS and LASSO regression. While LASSO regression produced smaller model sizes, NNLS had a faster runtime. We achieved fast and accurate results using both techniques. The LASSO regression L1 norm penalty shrinks abundance values toward zero resulting in successful sparse regression. Thus, LASSO regression performed well despite the large spectral library needed for unmixing complex mixtures. Its also interesting to note the computation time remained relatively low for the BMA approaches despite the ensambling of hundreds of models and regardless of the number of pixels in the ROI. This led us to ensambling sparse regression models with the BMA techniques to identify the ensambling strengths associated with sparse regression, in our next study. This approach may further enhance performance by leveraging sparsity-promoting properties.

3 Unmixing with Large Spectral Libraries for Identifying Soil and Mineral Chemistry

3.1 Introduction

In this chapter, we compare variations of two sparse regression techniques, focusing on the relationship between structure and chemistry of materials and the accuracy of the various models for identifying the correct mixture of materials present. Specifically, we examine LASSO regression and ElasticNet in contrast with variations of iterative feature selection, BMA, and BMA-Q — incorporating LASSO regression and ElasticNet as their base model. To evaluate the effectiveness of these methods, we consider the molecular composition similarities and differences of substances selected in the models compared to the ground truth.

3.2 Unmixing Techniques

LASSO regression and ElasticNet are frequently implemented regularization regression techniques. We incorporated a basic version of both techniques to identify their individual advantages as they both encourage sparse solutions. Python’s sklearn “LASSO” and “ElasticNet” packages were employed for the basic implementation.

For the iterative approaches, we independently developed (without the use of a Python package) two variations of DFS feature selection strategies. Our feature selection variations implemented LASSO regression and ElasticNet. We incorporated these variations to

simultaneously enforce sparsity and reduce modeling error.

LASSO and ElasticNet were also incorporated in a basic BMA implementation and BMA with quadratic terms, BMA-Q. We compared these sparse versions of BMA and BMA-Q to non sparse versions using NNLS. In total, we generated six different comparisons using BMA: BMA NNLS, BMA LASSO, BMA ElasticNet, BMA-Q NNLS, BMA-Q LASSO, and BMA-Q ElasticNet. The BMA-Q techniques included interactions of materials up to the 2nd order. We independently developed all of the BMA variations rather than using a Python package.

Table 3.1 presents the alpha parameters used for each method to achieve the unmixing results. Each technique is evaluated based on RMSE, model size, computation time, and detection percentage. Computation time refers to the average time required for each technique to unmix a single pixel within the ROI. Detection percentage indicates the proportion of inferred models in the ROI that successfully contain the target mineral. Unmixing performance may vary based on factors such as the target mineral, pixel location, and the composition of the mineral mixture. This comparison aims to determine the most effective techniques in certain scenarios.

	LASSO	ElasticNet	DFS LASSO	DFS ElasticNet	BMA NNLS	BMA LASSO	BMA ElasticNet	BMA-Q NNLS	BMA-Q LASSO	BMA-Q ElasticNet
Alpha	0.0004	0.001	0.0001	0.001	-	0.0001	0.0001	-	0.0001	0.001

Table 3.1: Presentation of alpha parameters for each unmixing technique. The parameter α controls the regularization strength for techniques that use regularization, while ‘-’ indicates methods where α is not applicable.

3.3 Physical-chemical Taxonomy

Every mineral is comprised of a chemical make up which dictates their physical appearance and texture. Mineralogist use this chemical structure to define and categorize these substances. Conveniently, these mineral types follow certain spectral patterns. We not only examined the mineral category of our target mineral, alunite and kaolinite, but we also identified the categories commonly detected with the target mineral type. Figure 3.1 dis-

plays six mineral types: sulfate, hydroxide, and inosilicate, phyllosilicate, sorosilicate, and cyclosilicate. The spectral library encompassed approximately twenty categories, but the mineral types referenced here frequently occur in the unmixing results. Each of the subplots in Figure 3.1 displays the pattern of a few of the minerals within a certain category from the spectral library.

3.4 Results

Category	Technique	RMSE Mean	Model Size	Runtime (s)	Detection
Regularization Approaches	LASSO	0.1310	4.3740	0.0050	0.9593
	ElasticNet	0.1389	7.0650	0.0033	0.9593
Iterative Approaches	DFS LASSO	0.0868	2.0894	1.0946	0.9268
	DFS ElasticNet	0.0857	2.1870	1.0772	0.9268
Ensembling Approaches	BMA NNLS	0.0519	2.7724	0.0487	0.8862
	BMA LASSO	0.0538	2.0732	0.6012	0.7398
	BMA ElasticNet	0.0525	2.1870	0.6168	0.7480
Nonlinear Approaches	BMAQ NNLS	0.2743	3.1382	0.0760	0.8455
	BMAQ LASSO	0.2452	2.4309	0.6928	0.5203
	BMAQ ElasticNet	0.2696	2.3008	0.6403	0.3333

(a) Metrics for unmixing techniques for the alunite ROI.

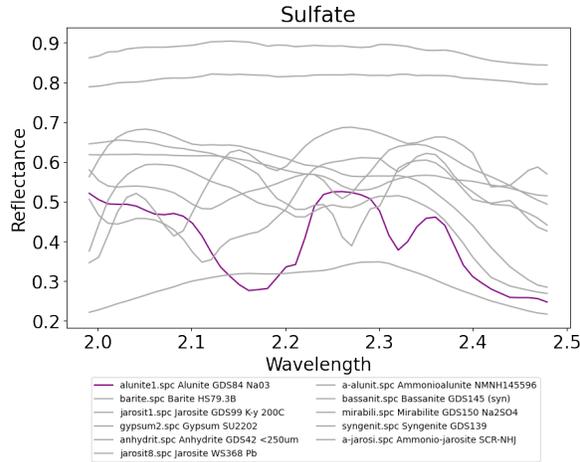
Category	Technique	RMSE Mean	Model Size	Runtime (s)	Detection
Regularization Approaches	LASSO	0.1348	6.8917	0.0143	0.9667
	ElasticNet	0.1384	12.5167	0.0099	1.0000
Iterative Approaches	DFS LASSO	0.0732	2.3167	1.3576	0.7583
	DFS ElasticNet	0.0733	2.4500	1.3540	0.7917
Ensembling Approaches	BMA NNLS	0.0630	3.6750	0.0730	0.6667
	BMA LASSO	0.0484	2.7083	0.6422	0.6833
	BMA ElasticNet	0.0493	2.6917	0.6552	0.6833
Nonlinear Approaches	BMAQ NNLS	0.3863	4.6250	0.1529	0.6167
	BMAQ LASSO	0.4923	2.3667	0.6293	0.3000
	BMAQ ElasticNet	0.4125	2.5417	0.6891	0.3000

(b) Metrics for unmixing techniques for the kaolinite ROI.

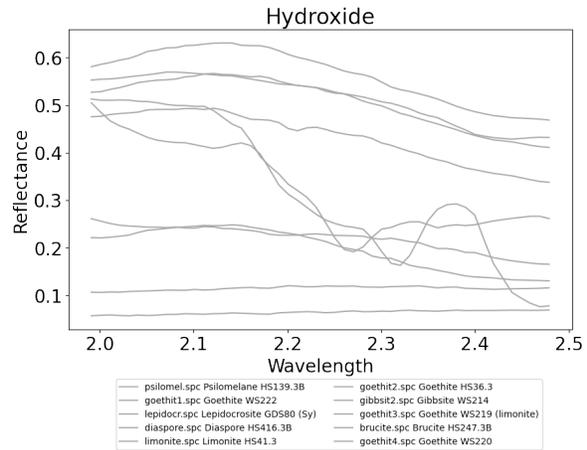
Table 3.2: Comparison of unmixing techniques with metrics including RMSE mean, model size, runtime, and detection rate for two datasets.

3.4.1 Unmixing Performance

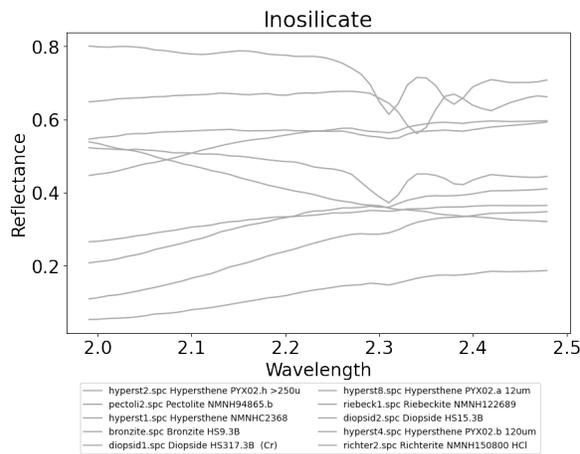
Table A.1 compares the unmixing performance of LASSO regression, ElasticNet, DFS using LASSO, DFS using ElasticNet, BMA using NNLS, BMA using LASSO, BMA using Elastic-



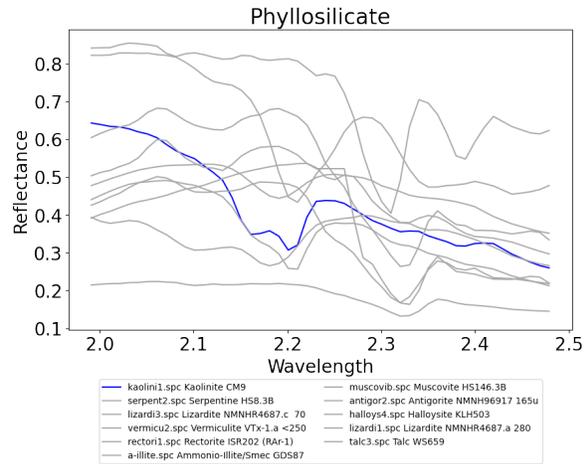
(a) Chemical Formula: SO_4



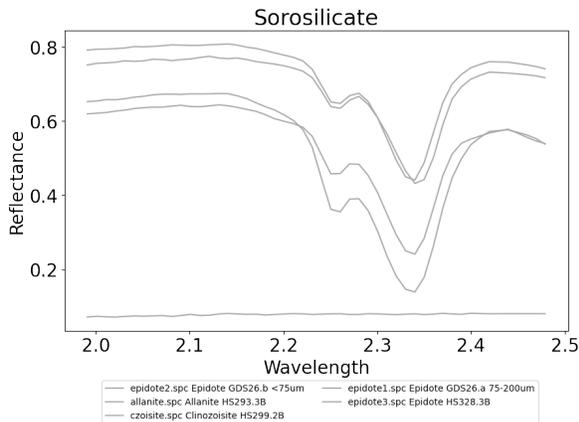
(b) Chemical Formula: OH



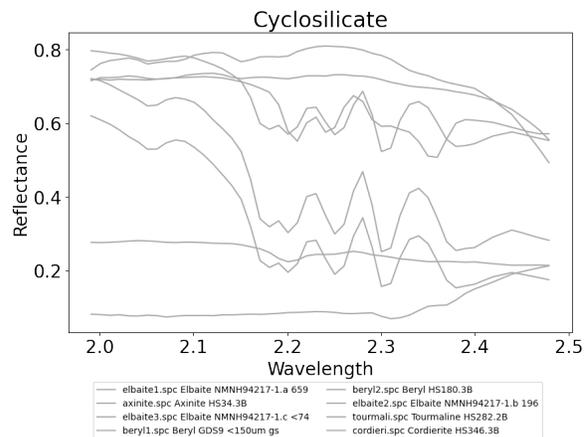
(c) Chemical Formula: SiO_3



(d) Chemical Formula: Si_2O_5

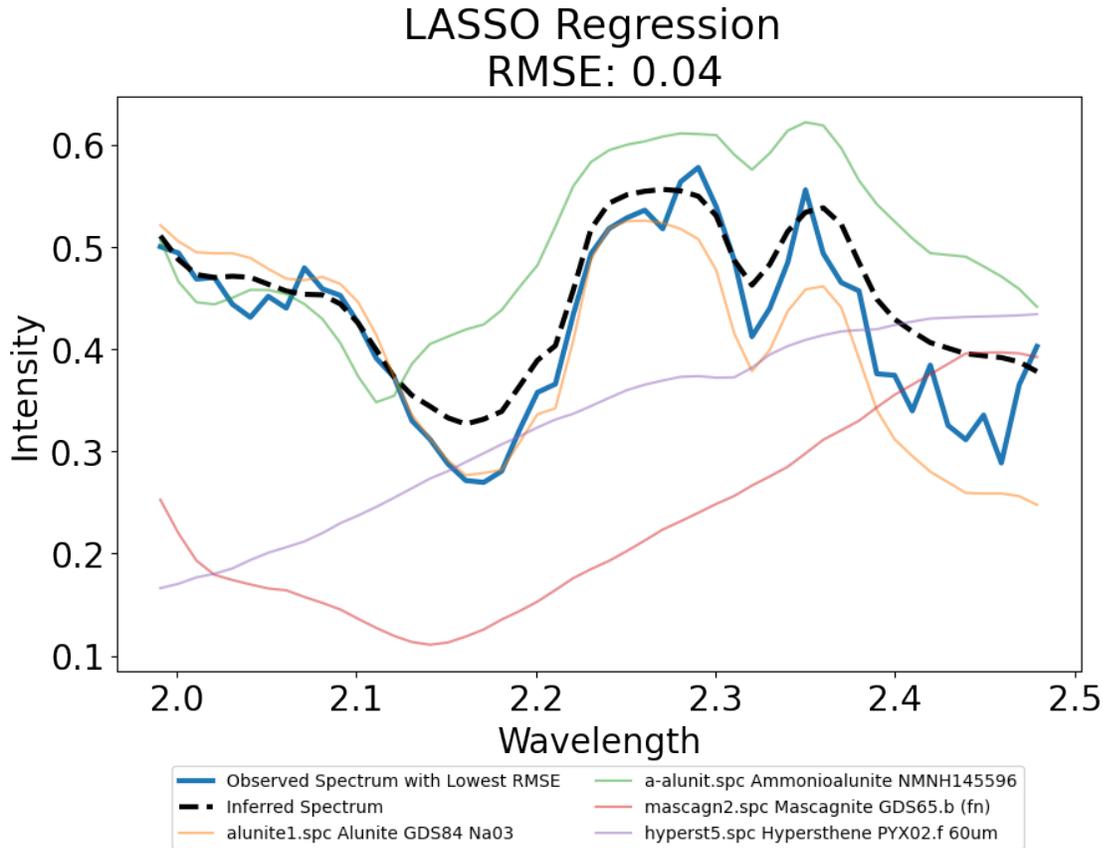


(e) Chemical Formula: Si_2O_7



(f) Chemical Formula: SiO_3

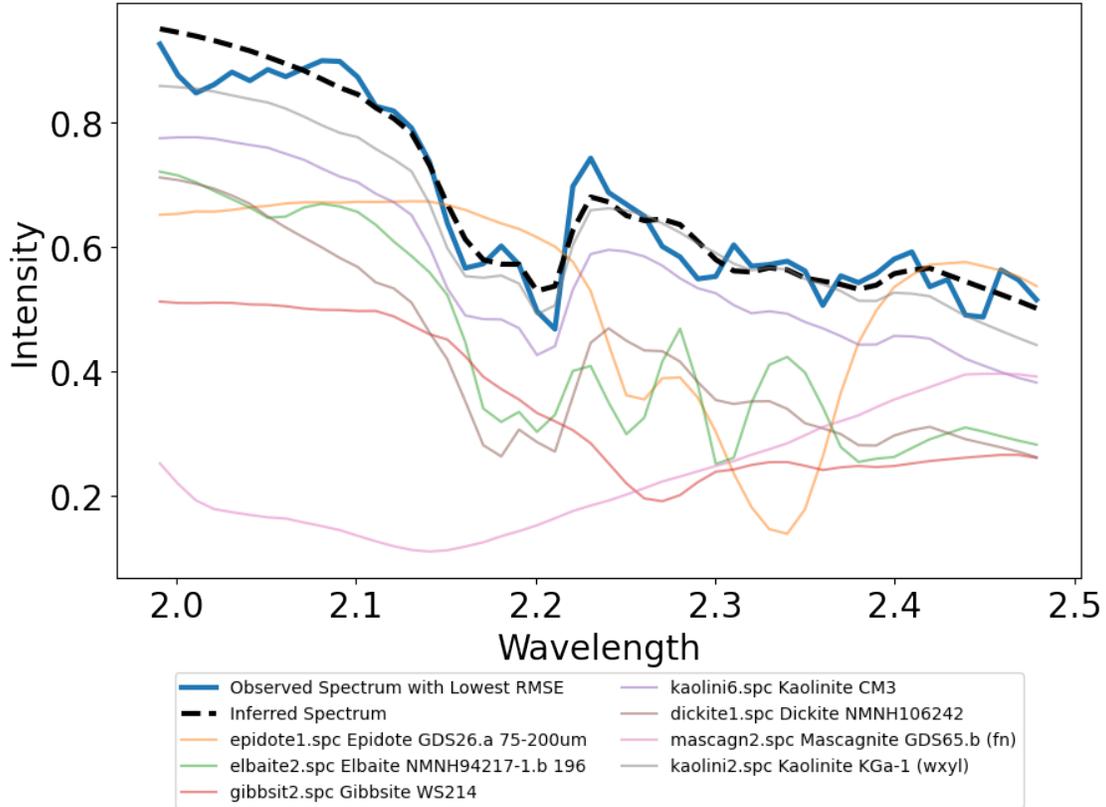
Figure 3.1: Display of the mineral types sulfate, hydroxide, inosilicate, phyllosilicate, sorosilicate, and cyclosilicate. Figure 3.1a shows the target mineral alunit highlighted in purple. Figure 3.1d shows the target mineral kaolinite highlighted in blue.



Name	Category	Formula	Abundance
alunit1.spc Alunite GDS84 Na03	Sulfate	$(\text{Na, K})\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.6641
a-alunit.spc Ammonioalunite NMNH145596	Sulfate	$(\text{NH}_4)\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.1950
mascagn2.spc Mascagnite GDS65.b (fn)	Sulfate	$(\text{NH}_4)_2\text{SO}_4$	0.1667
hyperst5.spc Hypersthene PYX02.f 60um	Inosilicate	$(\text{Mg, Fe}^{2+})_2\text{Si}_2\text{O}_6$	0.1430

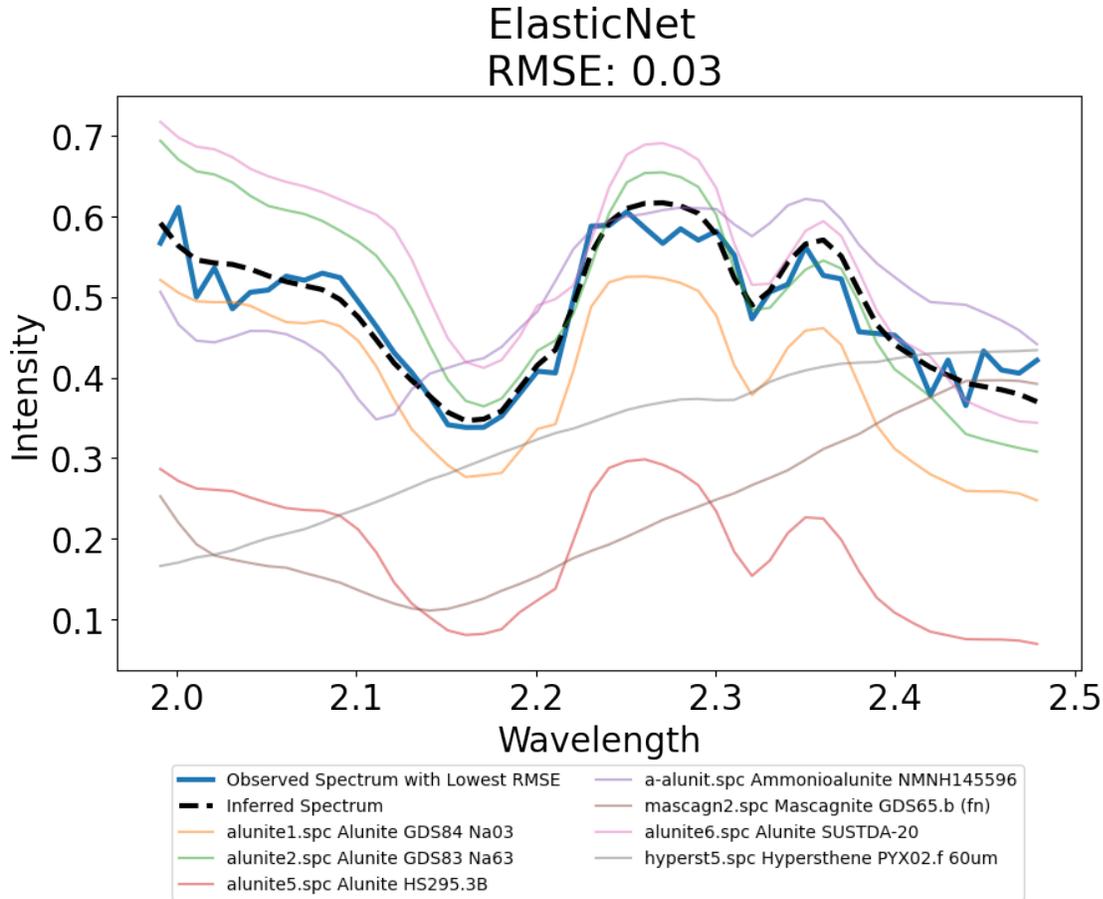
Figure 3.2: Alunite Table and Plot. Displays the inferred spectrum using LASSO regression with the most common minerals and their average abundances. Rows containing alunite are highlighted in yellow.

LASSO Regression
RMSE: 0.04



Name	Category	Formula	Abundance
epidote1.spc Epidote GDS26.a	Sorosilicate	$\text{Ca}_2(\text{Al,Fe}^{+3})_3(\text{SiO}_4)_3(\text{OH})$	0.0932
elbaite2.spc Elbaite NMNH94217-1.b	Cyclosilicate	$\text{Na}(\text{Li,Al})_3\text{Al}_6(\text{BO}_3)_3\text{Si}_6\text{O}_{18}(\text{OH})_4$	0.0966
gibbsit2.spc Gibbsite WS214	Hydroxide	$\text{Al}(\text{OH})_3$	0.1172
kaolini6.spc Kaolinite CM3	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.3362
dickite1.spc Dickite NMNH106242	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.2205
mascagn2.spc Mascagnite GDS65.b	Sulfate	$(\text{NH}_4)_2\text{SO}_4$	0.1171
kaolini2.spc Kaolinite KGa-1 (wxyl)	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.3645

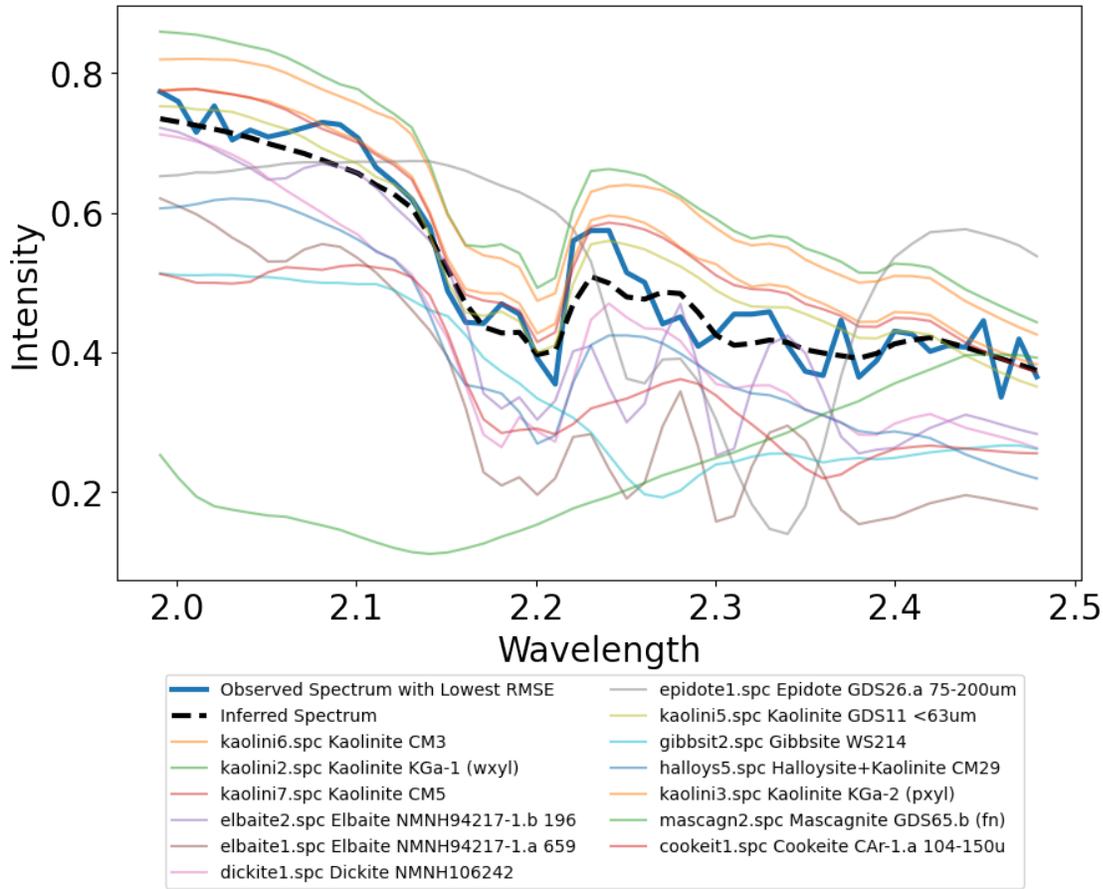
Figure 3.3: Kaolinite Table and Plot. Displays the inferred spectrum using LASSO regression with the most common minerals and their average abundances. Rows containing kaolinite are highlighted in cyan.



Name	Category	Formula	Abundance
alunite1.spc Alunite GDS84 Na03	Sulfate	$(\text{Na}, \text{K})\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.3596
alunite2.spc Alunite GDS83 Na63	Sulfate	$(\text{Na}, \text{K})\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.1706
alunite5.spc Alunite HS295.3B	Sulfate	$(\text{Na}, \text{K})\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.2280
a-alunit.spc Ammonioalunite NMNH145596	Sulfate	$(\text{NH}_4)\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.2285
mascagn2.spc Mascagnite GDS65.b (fn)	Sulfate	$(\text{NH}_4)_2\text{SO}_4$	0.1429
alunite6.spc Alunite SUSTDA-20	Sulfate	$(\text{Na}, \text{K})\text{Al}_3(\text{SO}_4)_2(\text{OH})_6$	0.0811
hyperst5.spc Hypersthene PYX02.f 60um	Inosilicate	$(\text{Mg}, \text{Fe}^{+2})_2\text{Si}_2\text{O}_6$	0.0640

Figure 3.4: Alunite Table and Plot. Displays the inferred spectrum using ElasticNet with the most common minerals and their average abundances. Rows containing alunite are highlighted in yellow.

ElasticNet RMSE: 0.04



Name	Category	Formula	Abundance
kaolini6.spc Kaolinite CM3	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.1087
kaolini2.spc Kaolinite KGa-1 (wxyz)	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.1046
kaolini7.spc Kaolinite CM5	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.1041
elbaite2.spc Elbaite NMNH94217-1.b 196	Cyclosilicate	$\text{Na}(\text{Li},\text{Al})_3\text{Al}_6(\text{BO}_3)_3\text{Si}_6\text{O}_{18}(\text{OH})_4$	0.0759
elbaite1.spc Elbaite NMNH94217-1.a 659	Cyclosilicate	$\text{Na}(\text{Li},\text{Al})_3\text{Al}_6(\text{BO}_3)_3\text{Si}_6\text{O}_{18}(\text{OH})_4$	0.0597
dickite1.spc Dickite NMNH106242	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.1696
epidote1.spc Epidote GDS26.a 75-200um	Sorosilicate	$\text{Ca}_2(\text{Al},\text{Fe}^{+3})_3(\text{SiO}_4)_3(\text{OH})$	0.0951
kaolini5.spc Kaolinite GDS11 <63um	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.0576
gibbsit2.spc Gibbsite WS214	Hydroxide	$\text{Al}(\text{OH})_3$	0.0673
halloys5.spc Halloysite+Kaolinite CM29	Phyllosilicate	Mixture of halloysite and kaolinite	0.0738
kaolini3.spc Kaolinite KGa-2 (pxyl)	Phyllosilicate	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$	0.0477
mascagn2.spc Mascagnite GDS65.b (fn)	Sulfate	$(\text{NH}_4)_2\text{SO}_4$	0.0885
cookeit1.spc Cookeite CAR-1.a 104-150u	Phyllosilicate	$\text{LiAl}_4(\text{Si}_3\text{Al})\text{O}_{10}(\text{OH})_8$	0.0416

Figure 3.5: Kaolinite Table and Plot. Displays the inferred spectrum using ElasticNet with the most common minerals and their average abundances. Rows containing kaolinite are highlighted in cyan.

Net, BMA-Q using NNLS, BMA-Q using LASSO, and BMA-Q using ElasticNet. The table shows the detection percentage, average model size, average computation time, and average RMSE. LASSO regression and ElasticNet performed better than all the other techniques in terms of computation time, and detection percentage. However, both techniques had the largest models as well as the highest modeling error of the linear approaches. ElasticNet achieved slightly better accuracy when unmixing pixels for the kaolinite ROI. Overall, the two regularization approaches had similar results ranging from 95-100% detection.

The iterative approaches had fairly high accuracy in detecting the target minerals and also achieved low error rates for both ROIs. DFS feature selection achieved a higher detection percentage with the alunite ROI (above 90% detection of the target mineral) than with the kaolinite ROI (detection rate between 75–80%). Despite achieving the second highest detection rates, the computation time for DFS was the highest across all methods implemented.

The detection percentage for BMA was lower compared to the sparse and iterative approaches. Furthermore, detection percentages for the BMA techniques were substantially lower with the kaolinite ROI than the alunite ROI. Also, it is interesting to note that the computation time for BMA NNLS was much faster than that for BMA LASSO and BMA ElasticNet. Even with the lower detection rates, the computed error for these ensambling techniques was lower than the sparse and iterative approaches. This means there probably exists an inferred model from both BMA LASSO and BMA ElasticNet that resembles a close fit to at least one of the pixels.

The nonlinear approaches achieved the lowest performance in terms of detection accuracy. Also, the detection percentages were spread over wider ranges — 33–84% with the alunite ROI and 30-62% with the kaolinite ROI. It’s also important to note, the average unmixing error was lower with the alunite ROI (RMSE Mean = 0.27, 0.25, 0.27) than with the kaolinite ROI (RMSE Mean = 0.40, 0.49, 0.41).

3.4.2 Feature Taxonomy

Figures 3.2 and 3.3 show the inferred models from LASSO regression and Figures 3.4 and 3.5 show ElasticNet for both ROIs. The figures display an inferred model generated from the most common minerals (technique specific) and their average estimated abundances. This inferred model is plotted with the closest fitting observed pixel spectrum and the constituent library spectra. For example, Figure 3.2 shows alunite, ammonioalunite, mascagnite and hypersthene as the most common minerals included in the inferred model using LASSO regression. Their average abundances using this technique are listed in the subtable of the figure. The table beneath the plot also shows each mineral's classification category and associated chemical formula.

3.5 Discussion

The results from this taxonomy indicate that many of the minerals included in the inferred models were from the same chemical category or categories with similar spectral features as the target mineral. Mascagnite was the common alternative sulfate incorporated in the model regardless of technique for the alunite ROI. Hypersthene – a inosilicate mineral – was the other frequently incorporated mineral.

Similarly, the models for the kaolinite ROI included minerals from the same phyllosilicate category or a mineral from a chemical category with similar features in specific bands. Dickite was a common phyllosilicate incorporated. Three other common mineral categories — for the kaolinite ROI — were cyclosilicate, sorosilicate and hydroxide. Models incorporated variations of elbaite from the cyclosilicate category. Epidote and gibbsite (respectively) were the most frequently incorporated sorosilicate and hydroxide.

3.6 Conclusion

In this study we compared the performance of 10 unmixing techniques. We measured performance based on the proportion of inferred spectra that included the target mineral, error metrics, model size and the unmixing computation time. The basic regularization approaches performed the best because they quickly detected the target mineral for almost every pixel. LASSO and ElasticNet had mildly better performance depending on the ROI, therefore we recommend implementing both methods for unmixing intimate mixtures.

A notable observation from the ElasticNet unmixing results was the detection of multiple variations of the target mineral. Both sparse techniques were able to positively detect alunite and kaolinite with high accuracy, but ElasticNet had several variation of the target material from the spectral library included in the models. ElasticNet involves both the L1 and L2 norm. Modifying the penalty term in ElasticNet caused the model to be less sparse than LASSO regression but sparse enough to allow only close variations. This inclusion of near-variations may reflect ElasticNet's ability to capture subtle spatial nuances and underlying phenomenology.

Intra-class variations were not the only variability observed from the taxonomy. We observed inter-class variability as well. Minerals from different chemical categories and their slight variations were often included. Certain mineral categories resembled the target mineral category within specific bands and therefore were included in the model to obtain a closer model fit. Furthermore, if the target was not included in the model, it was commonly replaced with a mineral from the same chemical category.

4 Rethinking Unmixing: A Comprehensive Evaluation of Algorithms

4.1 Introduction

In this chapter, we examine variations of OLS models stemming from highly effective approaches in spectral unmixing – sparse regression, iterative feature search strategies and Mathematical programming. These variations are compared to a novel unmixing approach called HySUDeB. We evaluated each approach’s performance by computing the average error and precision of each model. Additionally, we provide a taxonomy of the molecular structure of each mineral to derive further understanding into the detection of the target materials.

4.2 Experimental Design

The specific ROIs within the AVIRIS image are comprised of pixels from alunite, kaolinite, and montmorillonite locations. We used a graphical user interface, “Hyperspectralpy” [44] and the adaptive coherence estimator (ACE) detection algorithm to select the pixels within the ROIs. “Hyperspectralpy” is a Python package that provides enhanced analysis and visualization of hyperspectral images. The ACE algorithm selects target mineral spectra from the spectral library and calculates the average spectrum. A whitening transformation is then applied to the image spectra. Following the whitening process, an ACE score is computed depicting the strength of the target mineral’s presence. Figure 4.1 displays the

selected ROIs via the red colored regions. The top row in Figure 4.1 are the plots of the ACE detection score for the target mineral. An increased presence of alunite, kaolinite, and montmorillonite are indicated by the dark red. Consequently, we selected pixels with higher computed ACE scores (shown in the bottom row of Figure 4.1).

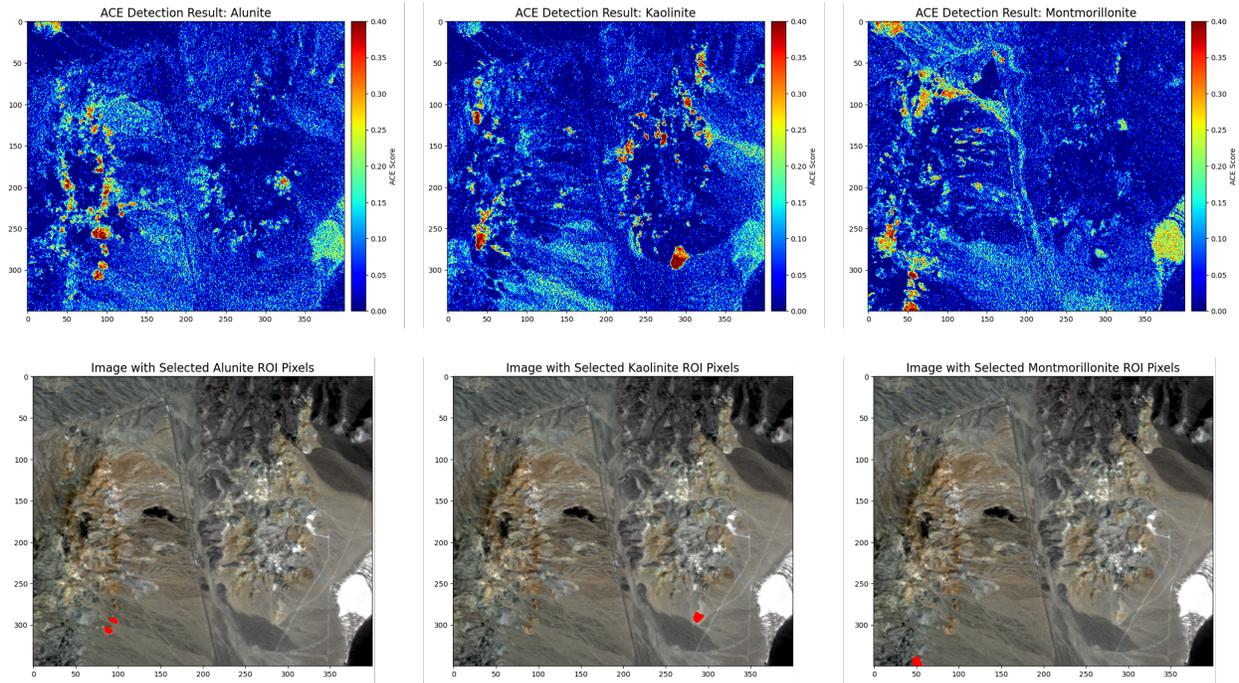


Figure 4.1: Display of the selected ROIs for three target minerals: alunite, kaolinite, and montmorillonite. The top row of plots show the regions where the target mineral shows higher abundance and the bottom row depicts the actual selected pixels from the locations with higher presence of the target mineral.

We then inferred models from the selected mixed pixels. Each inferred model represents the best fit of the observed spectrum. Subsequently, we recorded the category of each mineral from every model. Mineral categories are generally based on the physical-chemical or atomic bonding structure of the substance. These characteristics have both distinctive as well as similar properties, ultimately modifying their utility in material identification.

Figure 4.2 displays a few of the most common mineral categories of the spectra incorporated in each of the models. Each plot depicts the spectral pattern of minerals from specific categories. The physical patterns are indicative of their chemical structure. Sulfates have

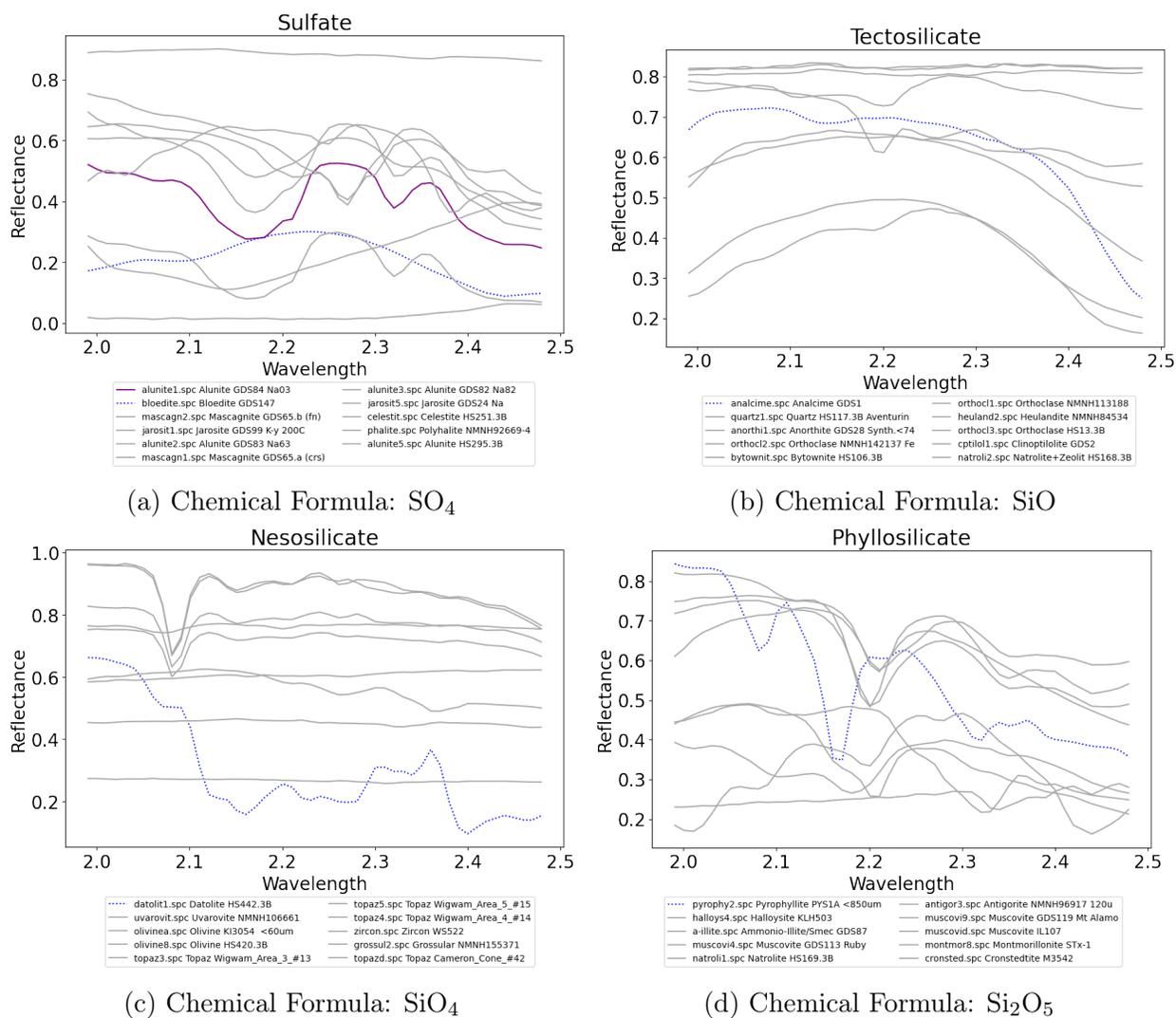


Figure 4.2: Comparison of the mineral types sulfate, tectosilicate, nesosilicate, and phyllosilicate. The purple spectra represents alunite and the dotted blue spectra are minerals that were often incorporated in the models.

a SO_4 chemical formula because they are comprised of one sulfur and four oxygen atoms. Tectosilicates have a SiO chemical formula which is indicative of their silicon and oxygen composition. Nesosilicates generally incorporate a silicon atom and four oxygen atoms, SiO_4 . Phyllosilicates have the chemical formula Si_2O_5 describing their composition of two silicon atoms and five oxygen atoms.

4.3 Statistical Analysis

4.3.1 LASSO Regression

LASSO regression was incorporated in this study because of its sparse solutions and known computational advantages. It responds to the disadvantages of OLS because it minimizes overfitting through noise reduction. The penalty parameter controlling sparsity was selected using cross validation (CV) ranging from 0.001 to 0.1 with 0.001 step size and $k=5$ fold. Additionally, we used the Python packages “Lasso” and “LassoCV” for implementation.

4.3.2 DFS Strategy

We implemented a DFS strategy because of its inherent feature selection capabilities. We developed the script for this technique and did not utilize a Python package. This approach can incur higher computational risk due to its iterative nature, but this allows for primary variable identification.

4.3.3 MINLP

The formulation of the MINLP within this paper has similarities to the models developed in Bourguignon et al. and Mhenni et al. with slight deviations [91], [99]. The following model is the mathematical program incorporated in this study.

Sets

I = set of all spectra in the library $i = [1, \dots, n] \in I$

J = set of all spectral bands $j = [1, \dots, m] \in J$

Parameters

y_j = observed pixel reflectance for band j

s_{ij} = reflected intensity for spectrum i in band j

M = number of spectral bands

B_i = max abundance associated with spectrum i

P = number of spectra included in the model

Decision Variables

x_i = decision to assign spectrum i to the model

a_i = abundance associated with spectrum i

Objective Function

$$\text{minimize } Z = \sqrt{\frac{1}{M} \sum_{j=1}^m \left(y_j - \sum_{i=1}^n (a_i s_{ij}) \right)^2}$$

Constraints

$$B_i x_i \geq a_i$$
$$\sum_i^n x_i \geq P$$

$$\forall x_i \in \{0, 1\}, \quad a_i \geq 0$$

Objective Explained

Minimize the root mean squared error between the observed spectrum and the inferred spectrum — the spectra incorporated in the model and their associated abundances.

Constraints Explained

1. Abundances can only be assigned to spectra included in the model and must adhere to the maximum material reflectance.
2. Model size constraint: The sum of the number of spectra assigned to the model must adhere to the model size parameter (P).
3. The decision to include a spectrum in the model (x_i) is binary. The abundance associated with a spectrum (a_i) must be positive.

4.3.4 HySUDeB

In this novel approach, we decorrelated the spectral bands as part of the unmixing process. This band decorrelation is also known as whitening. The OLS assumption is that the residuals follow a multivariate normal distribution with the mean centered at the pixel spectrum and covariance proportional to the identity matrix. However, this assumption may not always be appropriate, thus whitening becomes beneficial because we assume the covariance matrix of the spectral bands is not constant. Within the whitening process, we computed the covariance of our observed image. From the covariance matrix we derived the eigenvalues and eigenvectors. The library spectra and the observed spectrum were whitened by subtracting the mean and applying a transformation using the covariance matrix's eigenvalues and

eigenvectors [108]. Once band decorrelation was complete, we applied LASSO regression to the whitened spectra.

4.4 Technique Evaluation

We compared the unmixing performance of each technique based on computation time, error, detection accuracy, mean average precision at k , and the mineral category count detected. Detection accuracy was based on the number of target spectra inferred from the modeling approaches. We tallied the number of alunite, kaolinite, and montmorillonite detections achieved by each approach. Mean precision at k is often used for prediction of user recommendation in which recommendations are produced based on user preferences. We treated each model like a recommendation. Minerals were scored based on their assigned abundance proportion and the number of recommendations (i.e. number of minerals included in the model). Lastly, the mineral categories included were tallied across the models for each technique. For example, if a model incorporated two sulfate minerals, the category count was two and the calculated count was added to the total category count for the ROI.

4.5 Results

Figure 4.3 displays each technique's performance for the alunite ROI in terms of computation time and root mean squared error (RMSE). Though Figure 4.3 displays the result for the alunite ROI and the resulting plots for kaolinite and montmorillonite have different numeric values, the performance conclusion is the same for each of the ROIs in terms of this average RMSE vs. average computation time plot. LASSO regression had the lowest average computation time. MINLP had the lowest average residual error but far higher computation time than the other three techniques. DFS feature selection was not the highest performer in either area, but it shows to be the best overall for speed and minimal error. HySUDeB had similar computation time as LASSO regression, but because the units on the data for

HySUDeB are different than the units for the other methods, the computed RMSE was much higher.

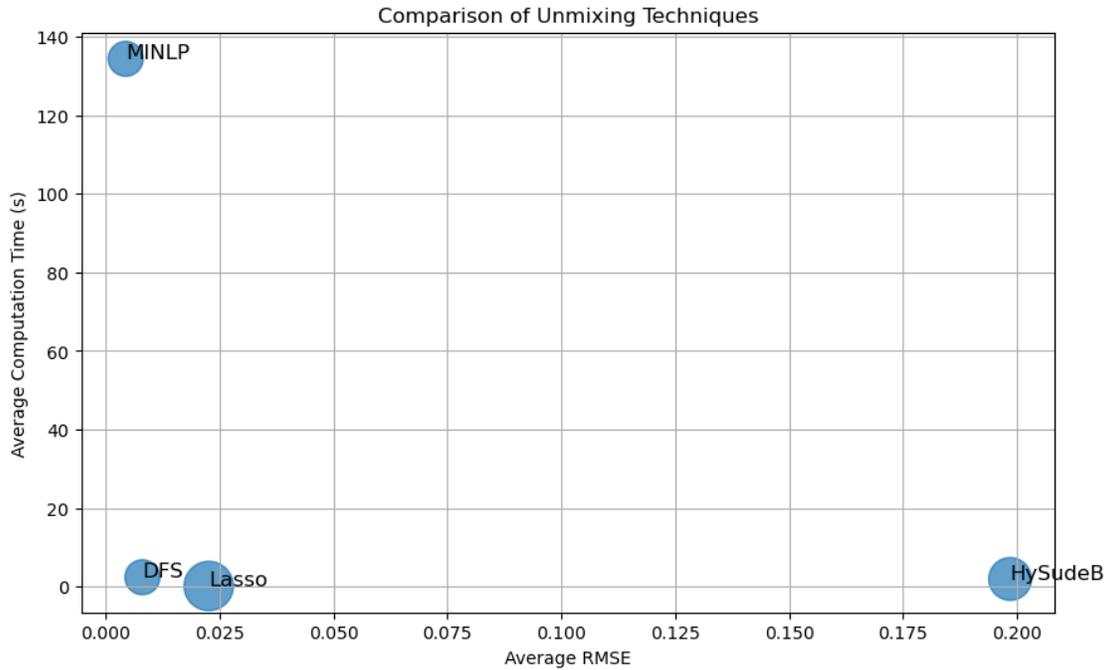


Figure 4.3: Comparison of the performance of each unmixing technique based on error and computation time.

The results for detection accuracy are shown in Table 4.1. For each of the ROIs, LASSO regression incorporated the highest number of target minerals across each of its models. MINLP had the second highest detection percentages. When target variables were incorporated in the solutions, MINLP only included the target mineral once, while other approaches incorporated variations of the target in many models.

Target Detection Percentage				
	Lasso	DFS	MINLP	HySudeB
Alunite	1.0	1.0	1.0	0.95
Kaolinite	1.0	0.65	1.0	0.94
Montmorillonite	0.85	0.56	0.77	0.22

Table 4.1: Detection percentages for target minerals (alunite, kaolinite, and montmorillonite) across four techniques: Lasso, DFS, MINLP, and HySudeB.

Table 4.2 displays the results in terms of detection precision. Precision at k is an indi-

cation of importance placed on the target mineral. Minerals included in each model were ranked based on model size and assigned abundance. Specifically, precision at k measures the number of times the target mineral was one of the top ordered materials. This precision table shows that HySUDeB performed the best with the alunite ROI, and MINLP performed the best with the other two regions, kaolinite and montmorillonite.

	Lasso	DFS	MINLP	HySudeB
Alunite	0.41	0.58	0.36	0.60
Kaolinite	0.39	0.31	0.51	0.46
Montmorillonite	0.14	0.14	0.18	0.03

Table 4.2: Average mean precision at k for minerals (alunite, kaolinite, and montmorillonite) across four techniques: Lasso, DFS, MINLP, and HySudeB.

In addition to measuring performance, we identified the mineral categories incorporated in each modeling approach. The most commonly included categories regardless of technique were sulfates, tectosilicates, nesosilicates, carbonates, and phyllosilicates. Figure 4.4 shows the count for all the included mineral categories across all three ROIs. The categories shown in Figure 4.4 are not an exhaustive list of the mineral types found in the spectral library. Figure 4.5 shows a more extensive list of the primary library categories as well as the counts based on unmixing approach. LASSO regression showed to have the most difference in mineral inclusion but the lowest diversity in mineral category for each ROI. For example, in Figure 4.5, Hydroxide minerals were included in all the other approaches except LASSO. DFS feature selection and HySUDeB had similar spreads of included mineral categories. Both approaches show a definitive gap between the highest counts and second highest counts. Even though LASSO regression included the highest amount of sulfates in the models for the alunite ROI — presumably because alunite is a sulfate — LASSO regression included nearly as many phyllosilicates. HySUDeB was a close second to LASSO regression in terms of including sulfates in the model for the alunite ROI, however unlike LASSO regression there is a distinct difference between the count for sulfates and phyllosilicates. HySUDeB shows this pattern of category inclusion with the kaolinite and montmorillonite ROIs as well.

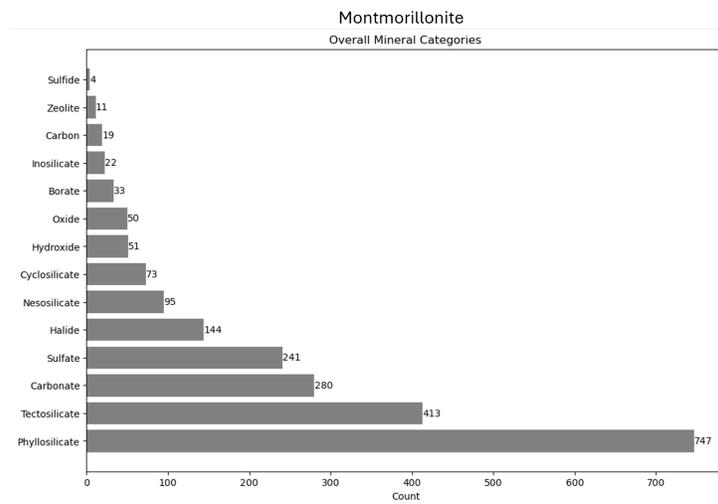
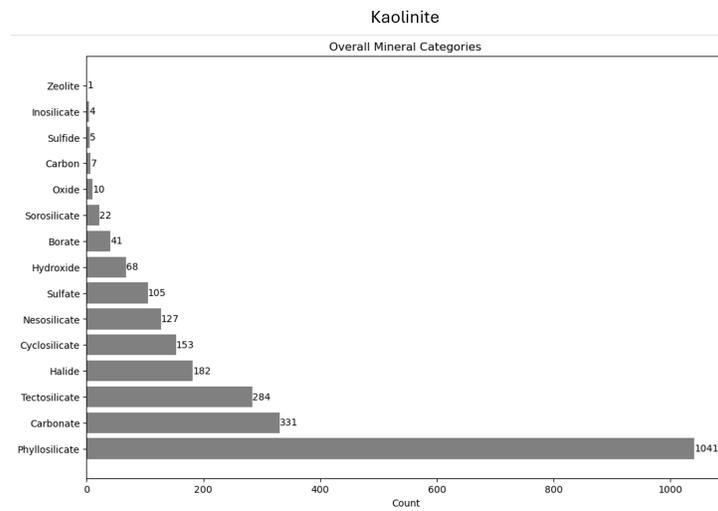
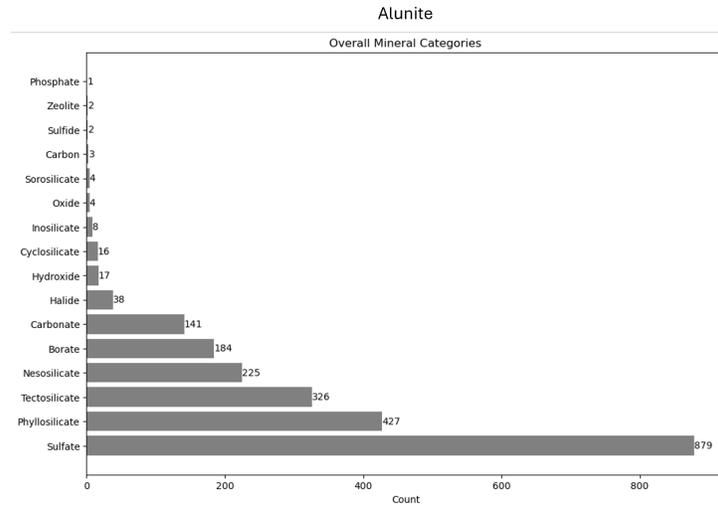


Figure 4.4: Display of the count of the included minerals across all the implemented approaches.

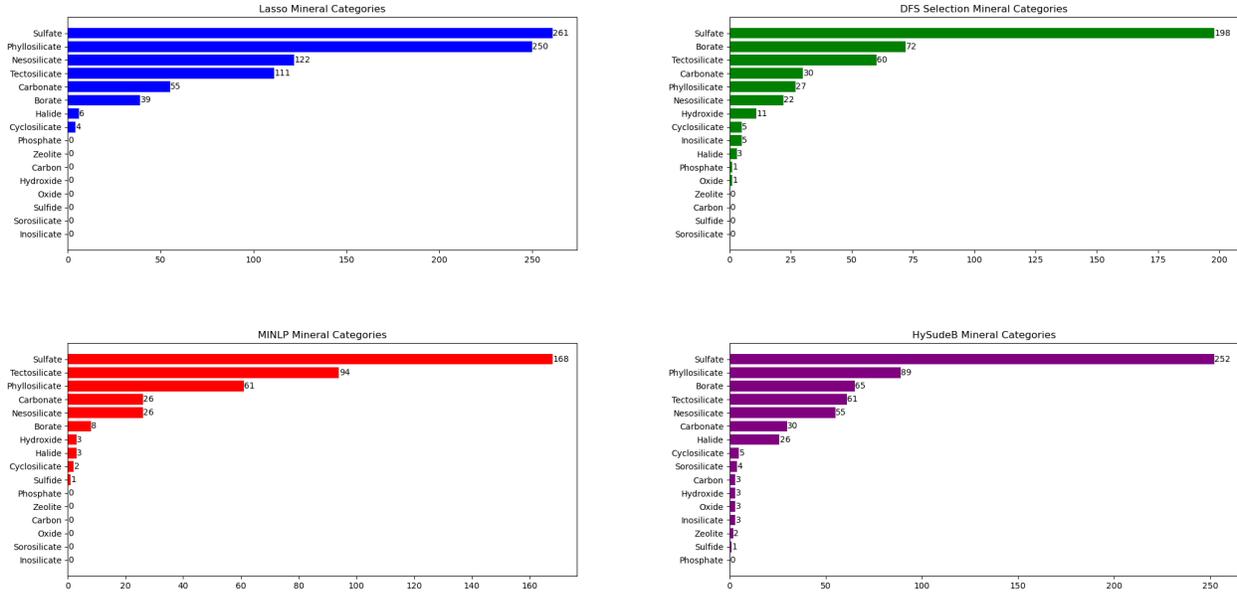


Figure 4.5: Display of the count of the mineral from the library that were included in the models for the alunite ROI.

4.6 Discussion

This study compared four unmixing techniques using four aspects: computation time, error minimization, detection accuracy, and mean average precision at k . Our results indicate that the employed techniques dominate in specific tasks and goals. No one method was universally superior, but each is highly effective depending on the aim of the study. LASSO regression was the fastest unmixing technique. MINLP had the lowest minimal error and DFS feature selection achieved the best balance across performance comparison. HySUDeB performed well in detecting the target mineral category and thus features associated with those categories.

The goal of this study was primary mineral detection. The established penalty parameter for LASSO regression allowed the model to become less sparse. This reduced sparsity ensured: 1) the target mineral was included in the model at least 85% of the time and 2) LASSO regression obtained the highest cumulative target mineral count. However, this also caused the model sizes to be larger and ultimately lowered the mean average precision at k .

HySUDeB performed well in terms of feature detection of the target mineral — as did DFS feature selection. For each of the ROIs, the chemical category of the target mineral was definitive because HySUDeB addresses the OLS homoscedasticity failures. In this approach, we decorrelated spectral bands, effectively whitening the data. By transforming the spectral data into a space where the bands were uncorrelated, we stabilized the variance of the residuals and improved the reliability of our model estimates.

The decorrelation process in HySUDeB enhanced the precision scoring of specific minerals, such as alunite. Band decorrelation ensured evenly distributed error, enabling more precise identification of mineral signatures in the presence of varying noise levels across spectral bands. Computation time was not affected with the implementation of whitening, but the error had different units and thus comparing RMSE was not meaningful. Future work could focus on refining techniques like HySUDeB to decrease the residual error while also maintaining the computation and detection capabilities.

5 Conclusion and Future Work

This dissertation presents theoretical and practical advancements in the field of Hyperspectral Image Analysis, with particular emphasis on the spectral unmixing domain. Our focus centered on unmixing intimate mixtures with the dual objective of advancing analysis across all measures of mixture compositional complexity. By achieving accurate unmixing for mineral deposits within soil and detecting target minerals with satisfactory precision, we can confidently unmix pixel spectra and evaluate technique performance in less complex compositions. A core challenge in unmixing intimate mixtures lies within spectral variability and consequently, the necessity of extensive spectral libraries for material comparison resulting in an ill-posed inversion problem. Several techniques for intimate unmixing have been proposed, but few studies systematically evaluate unmixing techniques.

5.1 Summary of Contributions

In Chapter 2, we unmixed two ROIs (alunite and kaolinite) and compared the performance of unconstrained OLS with partially constrained least squares, regularization, iterative search optimization, and two versions of BMA (i.e. linear and nonlinear ensembling). We presented evaluation metrics as a standard framework for comparing unmixing techniques. The metrics included in this comparison were model size, computation time, average error, and detection percentage. Through this study, we showed that OLS and regularization using ridge regression were poor performers for unmixing because they produce impractical mixture models, but regularization in the form of model size penalties — or parameter constraints such as

NNLS and LASSO regression — resulted in high detection percentages.

Chapter 3 focused on evaluating variations of sparse unmixing for the alunite and kaolinite ROIs. Because Ridge and LASSO regression performed so differently, we explored their combination through unmixing with ElasticNet. Basic LASSO regression, basic ElasticNet as well as feature search optimization and ensembling versions of LASSO regression and ElasticNet were compared in terms of performance. Additionally, we examined the physical-chemical structure of the minerals included in the models. The sparse approaches achieved the highest performance for both ROIs.

In Chapter 4, we compared LASSO regression, DFS feature selection and mathematical programming to a novel unmixing approach in which we decorrelated the spectral bands as part of the unmixing process. Sparse regression and DFS feature selection were incorporated in this study because of their high performance achieved in the previous studies. Mathematical programming has been known to achieve high levels of accuracy in other studies — with the acceptance of high computation time — therefore, we included a MINLP to assess any other unmixing strengths. We developed HySUDeB to address the OLS homoscedastic misalignment with spectral unmixing. By whitening the spectra, we removed band correlation. The performance evaluation metrics incorporated in the study were average computation time, average error, target detection percentage, average mean precision at k, and the mineral category count. Within this study, we continued our exploration of the physical chemical patterns associated with the target mineral. No spectral unmixing study has incorporated average mean precision at k as a metric for technique evaluation. However, average mean precision at k evaluates the technique based on the weight (or abundance) they assign to the target.

Model size, computation time, average error, target detection percentage, average mean precision at k, and the mineral category count were included in this dissertation because of their evaluation strengths. OLS included the entire spectral library, so we decided to include model size as a metric. Ensembling was an approach initiated in the first study,

Chapter 2, and incorporated in the follow-on research thrust, Chapter 3. Because thousands of models were combined for unmixing through the ensembling approaches, we assumed time complexity for unmixing may increase. Thus, we used computation time as another metric. We implemented RMSE as an evaluation metric because it is a common metric for regression models. However, through this dissertation, we noticed frequent scenarios in which RMSE was low and produced an inferred model with a close fit to the observed spectrum, but lacked the target mineral. Our main goal was to identify the target mineral thus, we incorporated percent detection as an additional metric for technique evaluation. Average mean precision at k takes percent detection a step further by evaluating the techniques based on detection of the target and also the abundance assigned to the target. Furthermore, average mean precision at k incorporates some of the other metric strengths such as model size. Larger models resulted in lowered mean precision at k values. Finally, we included mineral category count as a metric specifically to assess the detection of the physical-chemical patterns.

5.2 Future Work

Unmixing vegetation as well as granular man-made substances would serve as a powerful continuation of this work particularly in developing the chemical-physical chemical taxonomy. Vegetation or synthetic material detection and identification would aid in disaster relief and military use cases. There is an abundance of research that can be accomplished in understanding the molecular and chemical bonds for these material types.

Another useful continuation of this work would be determining relevant spectral bands as it relates to the physical chemical design of material. Research can be developed identifying the exact range of wavelengths on the electromagnetic spectrum to prioritize unmixing materials in specific chemical categories. We categorized spectral patterns associated with chemical categories of material, but band isolation for material types would be a valuable research thrust.

5.3 Concluding Remarks

Mineral deposits are a well studied and readily available dataset. We presented our research as an application into mineral detection because the complexity of this mixture composition forced the resolution of an important and prevalent problem in hyperspectral analysis — spectral unmixing with a large reference library. Our results demonstrate the performance and nuanced capabilities of unmixing techniques when applied to a large real world spectral library. The findings within this body of work are instrumental for future research and broader advancements in various material identification and detection applications.

This dissertation compared a wealth of unmixing techniques as well as provided an a evaluation framework robust for a variety of mixture compositional complexities. We determined the approaches with high performance based on various objectives for multiple targets. The analysis of physical-chemical structures shows that the models included the target mineral, replaced the target mineral with a mineral from the same category, or replaced the target mineral with a mineral possessing similar spectral pattern features.

We hope this body of work serves the field of Hyperspectral Image Analysis well and makes a small but useful contribution to the field of Data Science.

References

- [1] D. Manolakis, D. Marden, G. A. Shaw, *et al.*, “Hyperspectral image processing for automatic target detection applications,” *Lincoln laboratory journal*, vol. 14, no. 1, pp. 79–116, 2003.
- [2] P. Ghamisi *et al.*, “Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, 2017.
- [3] J. M. Bioucas-Dias *et al.*, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geoscience and remote sensing magazine*, vol. 1, no. 2, pp. 6–36, 2013.
- [4] R. A. Borsoi *et al.*, “Spectral variability in hyperspectral data unmixing: A comprehensive review,” *IEEE geoscience and remote sensing magazine*, vol. 9, no. 4, pp. 223–270, 2021.
- [5] B. Rasti *et al.*, “Image processing and machine learning for hyperspectral unmixing: An overview and the hysupp python package,” *arXiv preprint arXiv:2308.09375*, 2023.
- [6] R. N. Clark *et al.*, *Spectroscopy of rocks and minerals, and principles of spectroscopy*, 2020.
- [7] J. W. Chapman *et al.*, “Spectral and radiometric calibration of the next generation airborne visible infrared spectrometer (aviris-ng),” *Remote Sensing*, vol. 11, no. 18, p. 2129, 2019.

- [8] F. F. Sabins Jr and J. M. Ellis, *Remote sensing: Principles, interpretation, and applications*. Waveland Press, 2020.
- [9] S. Aggarwal, “Principles of remote sensing,” *Satellite remote sensing and GIS applications in agricultural meteorology*, vol. 23, no. 2, pp. 23–28, 2004.
- [10] E. J. Milton *et al.*, “Progress in field spectroscopy,” *Remote Sensing of Environment*, vol. 113, S92–S109, 2009.
- [11] R. R. Navalgund, V. Jayaraman, and P. Roy, “Remote sensing applications: An overview,” *current science*, pp. 1747–1766, 2007.
- [12] V. Lodhi, D. Chakravarty, and P. Mitra, “Hyperspectral imaging for earth observation: Platforms and instruments,” *Journal of the Indian Institute of Science*, vol. 98, pp. 429–443, 2018.
- [13] B. Lu *et al.*, “Recent advances of hyperspectral imaging technology and applications in agriculture,” *Remote Sensing*, vol. 12, no. 16, p. 2659, 2020.
- [14] J. B. Campbell and R. H. Wynne, *Introduction to remote sensing*. Guilford press, 2011.
- [15] C. J. Cohen, “Early history of remote sensing,” in *Proceedings 29th Applied Imagery Pattern Recognition Workshop*, IEEE Computer Society, 2000, pp. 3–3.
- [16] J. R. Jensen, *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India, 2009.
- [17] A. P. Cracknell, *The development of remote sensing in the last 40 years*, 2018.
- [18] S.-E. Qian, “Hyperspectral satellites, evolution, and development history,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7032–7056, 2021.
- [19] D. Landgrebe, “The evolution of landsat data analysis,” *Photogrammetric engineering and remote sensing*, vol. 63, no. 7, pp. 859–867, 1997.

- [20] D. Landgrebe, “Information extraction principles and methods for multispectral and hyperspectral image data,” in *Information processing for remote sensing*, World Scientific, 1999, pp. 3–37.
- [21] D. Landgrebe, “Hyperspectral image data analysis as a high dimensional signal processing problem,” *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 17–28, 2002.
- [22] D. Landgrebe, “Hyperspectral image data analysis,” *IEEE Signal processing magazine*, vol. 19, no. 1, pp. 17–28, 2002.
- [23] F. M. of Economic Affairs and C. Action, *Enmap successfully launched!!!* Apr. 2022. URL: <https://www.enmap.org/news/2022-04-04/>.
- [24] J. Werdell and D. Casey, *Nasa pace - home*, Mar. 2024. URL: <https://pace.gsfc.nasa.gov/>.
- [25] A. Wibisono, *Jupiter icy moons explorer (juice) mission: All you need to know*, Apr. 2023. URL: <https://www.rmg.co.uk/stories/blog/astronomy/jupiter-icy-moons-explorer-juice-mission>.
- [26] J. Zwinkels, “Light, electromagnetic spectrum,” *Encyclopedia of Color Science and Technology*, vol. 8071, pp. 1–8, 2015.
- [27] B. Basner, *Statistical machine learning for spectral image processing*, Unpublished manuscript, August 22, 2024, 2024.
- [28] M. Sivakumar, *Satellite remote sensing and gis applications in agricultural meteorology*, 2003.
- [29] R. Collier, “Transmission lines and photons,” in *Transmission Lines: Equivalent Circuits, Electromagnetic Theory, and Photons*, Accessed on October 30, 2024, Cambridge University Press, 2013, pp. 1–62. URL: <https://ebookcentral-proquest-com.proxy1.library.virginia.edu/lib/uva/detail.action?docID=1113081>.
- [30] G. Butcher, *Tour of the electromagnetic spectrum*. Government Printing Office, 2016.

- [31] G. A. Shaw and H. K. Burke, "Spectral imaging for remote sensing," *Lincoln laboratory journal*, vol. 14, no. 1, pp. 3–28, 2003.
- [32] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [33] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [34] J. Plaza *et al.*, "On endmember identification in hyperspectral images without pure pixels: A comparison of algorithms," *Journal of Mathematical Imaging and Vision*, vol. 42, pp. 163–175, 2012.
- [35] B. Basener, "Target identification and bayesian model averaging with probabilistic hierarchical factor probabilities," in *2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, IEEE, 2022, pp. 1–7.
- [36] W. Fan *et al.*, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, 2009.
- [37] N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln laboratory journal*, vol. 14, no. 1, pp. 55–78, 2003.
- [38] G. A. Swayze *et al.*, "Mapping advanced argillic alteration at cuprite, nevada, using imaging spectroscopy," *Economic geology*, vol. 109, no. 5, pp. 1179–1221, 2014.
- [39] R. O. Green *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris)," *Remote sensing of environment*, vol. 65, no. 3, pp. 227–248, 1998.
- [40] D. G. Goodenough *et al.*, "Comparison of aviris and aisa airborne hyperspectral sensing for above-ground forest carbon mapping," in *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, vol. 2, 2008, pp. II–129.

- [41] S. K. Meerdink *et al.*, “The ecostress spectral library version 1.0,” *Remote Sensing of Environment*, vol. 230, p. 111–196, 2019.
- [42] R. F. Kokaly *et al.*, “Usgs spectral library version 7,” US Geological Survey, Tech. Rep., 2017.
- [43] A. M. Baldridge *et al.*, “The aster spectral library version 2.0,” *Remote sensing of environment*, vol. 113, no. 4, pp. 711–715, 2009.
- [44] B. Basener, *Hyperspectralpy*, Nov. 2022. URL: <https://pypi.org/project/hyperspectralpy/>.
- [45] R. Clark *et al.*, “Usgs digital spectral library splib06a: Us geological survey, digital data series 231, 2007,” URL <http://speclab.cr.usgs.gov/spectral.lib06/ds231/index.html>, 2007.
- [46] D. Nash and J. Conel, “Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite,” *Journal of Geophysical Research*, vol. 79, no. 11, pp. 1615–1621, 1974.
- [47] B. Hapke, *Theory of reflectance and emittance spectroscopy*. Cambridge university press, 2012.
- [48] E. Ibarrola-Ulzurrun *et al.*, “Hyperspectral classification through unmixing abundance maps addressing spectral variability,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4775–4788, 2019.
- [49] M. Parente and A. Plaza, “Survey of geometric and statistical unmixing algorithms for hyperspectral images,” in *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, IEEE, 2010, pp. 1–4.
- [50] N. Dobigeon *et al.*, “Nonlinear unmixing of hyperspectral images: Models and algorithms,” *IEEE Signal processing magazine*, vol. 31, no. 1, pp. 82–94, 2013.

- [51] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *IEEE transactions on geoscience and remote sensing*, vol. 39, no. 7, pp. 1392–1409, 2001.
- [52] N. Keshava *et al.*, "Algorithm taxonomy for hyperspectral unmixing," in *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*, SPIE, vol. 4049, 2000, pp. 42–63.
- [53] I. Dópido *et al.*, "A quantitative and comparative assessment of unmixing-based feature extraction techniques for hyperspectral image classification," *IEEE Journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 421–435, 2012.
- [54] J. Preston and W. Basener, "Modeling uncertainty in hyperspectral image classification using neural networks with bayesian monte carlo dropout," in *2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2023, pp. 1–6. DOI: [10.1109/WHISPERS61460.2023.10430587](https://doi.org/10.1109/WHISPERS61460.2023.10430587).
- [55] J.-P. Combe *et al.*, "Analysis of omega/mars express data hyperspectral data using a multiple-endmember linear spectral unmixing model (melsum): Methodology and first results," *Planetary and Space Science*, vol. 56, no. 7, pp. 951–975, 2008.
- [56] R. Heylen and P. Gader, "Nonlinear spectral unmixing with a linear mixture of intimate mixtures model," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 7, pp. 1195–1199, 2013.
- [57] P. E. Johnson *et al.*, "A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures," *Journal of Geophysical Research: Solid Earth*, vol. 88, no. B4, pp. 3557–3561, 1983.
- [58] J. F. Mustard and C. M. Pieters, "Quantitative abundance estimates from bidirectional reflectance measurements," *Journal of Geophysical Research: Solid Earth*, vol. 92, no. B4, E617–E626, 1987.

- [59] B. Celik, “Qlsu (qgis linear spectral unmixing) plugin: An open source linear spectral unmixing tool for hyperspectral & multispectral remote sensing imagery,” *Environmental Modelling & Software*, vol. 168, p. 105782, 2023.
- [60] L. I. Jiménez and A. Plaza, “Hypermix: An open-source tool for fast spectral unmixing on graphics processing units,” *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1883–1887, 2015.
- [61] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, “Sparse unmixing of hyperspectral data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.
- [62] F. Schmidt, M. Legendre, and S. Le Mouëlic, “Minerals detection for hyperspectral images using adapted linear unmixing: Linmin,” *Icarus*, vol. 237, pp. 61–74, 2014.
- [63] K. E. Themelis *et al.*, “On the unmixing of mex/omega hyperspectral data,” *Planetary and Space Science*, vol. 68, no. 1, pp. 34–41, 2012.
- [64] A. L. Burton, “Ols (linear) regression,” *The encyclopedia of research methods in criminology and Criminal Justice*, vol. 2, pp. 509–514, 2021.
- [65] C. Dismuke and R. Lindrooth, “Ordinary least squares,” *Methods and designs for outcomes research*, vol. 93, no. 1, pp. 93–104, 2006.
- [66] M. A. Poole and P. N. O’Farrell, “The assumptions of the linear regression model,” *Transactions of the Institute of British Geographers*, pp. 145–158, 1971.
- [67] M. Tranmer and M. Elliot, “Multiple linear regression,” *The Cathie Marsh Centre for Census and Survey Research (CCSR)*, vol. 5, no. 5, pp. 1–5, 2008.
- [68] R. Wang, J. Du, and C. Ma, “Covariance matrix functions of isotropic vector random fields,” *Communications in Statistics-Theory and Methods*, vol. 43, no. 10-12, pp. 2081–2093, 2014.

- [69] C.-I. Chang, *Hyperspectral data processing: algorithm design and analysis*. John Wiley & Sons, 2013.
- [70] C.-I. Chang and K.-H. Liu, “Progressive band selection of spectral unmixing for hyperspectral imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2002–2017, 2013.
- [71] M. E. Winter, P. G. Lucey, and D. Steutel, “Examining hyperspectral unmixing error reduction due to stepwise unmixing,” in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX*, SPIE, vol. 5093, 2003, pp. 380–389.
- [72] D. C. Heinz *et al.*, “Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery,” *IEEE transactions on geoscience and remote sensing*, vol. 39, no. 3, pp. 529–545, 2001.
- [73] R. Heylen, D. Burazerovic, and P. Scheunders, “Fully constrained least squares spectral unmixing by simplex projection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4112–4122, 2011.
- [74] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, “A data dependent multiscale model for hyperspectral unmixing with spectral variability,” *IEEE Transactions on Image Processing*, vol. 29, pp. 3638–3651, 2020.
- [75] C. Loughlin *et al.*, “Efficient hyperspectral target detection and identification with large spectral libraries,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6019–6028, 2020.
- [76] J. K. Mitchell, K. Soga, *et al.*, *Fundamentals of soil behavior*. John Wiley & Sons New York, 2005, vol. 3.
- [77] D. d. Richter and D. H. Yaalon, ““the changing model of soil” revisited,” *Soil Science Society of America Journal*, vol. 76, no. 3, pp. 766–778, 2012.

- [78] R. V. Rossel *et al.*, “A global spectral library to characterize the world’s soil,” *Earth-Science Reviews*, vol. 155, pp. 198–230, 2016.
- [79] J. O. Ogutu, T. Schulz-Streeck, and H.-P. Piepho, “Genomic selection using regularized linear regression models: Ridge regression, lasso, elastic net and their extensions,” in *BMC proceedings*, Springer, vol. 6, 2012, pp. 1–6.
- [80] A. Bedoui and N. A. Lazar, “Bayesian empirical likelihood for ridge and lasso regressions,” *Computational Statistics & Data Analysis*, vol. 145, p. 106917, 2020.
- [81] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, “Collaborative sparse regression for hyperspectral unmixing,” *IEEE Transactions on geoscience and remote sensing*, vol. 52, no. 1, pp. 341–354, 2013.
- [82] R. M. Bastiaan, D. T. Salaki, and D. Hatidja, “Comparing the performance of prediction model of ridge and elastic net in correlated dataset,” in *Operations Research: International Conference Series*, vol. 3, 2022, pp. 8–13.
- [83] J. Li *et al.*, “Local spectral similarity preserving regularized robust sparse hyperspectral unmixing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 7756–7769, 2019.
- [84] J. Bioucas-Dias and M. Figueiredo, “An alternating direction algorithm for collaborative hierarchical sparse regularization,” *IST, Lisbon, Portugal, Tech. Rep.*, vol. 1, 2011.
- [85] S. Sun *et al.*, “Hyperspectral-multispectral image fusion using subspace decomposition and elastic net regularization,” *International Journal of Remote Sensing*, vol. 45, no. 12, pp. 3962–3991, 2024.
- [86] P. Ruengvirayudh and G. P. Brooks, “Comparing stepwise regression models to the best-subsets models, or, the art of stepwise,” *General linear model journal*, 2016.

- [87] A. M. Olusegun, H. G. Dikko, and S. U. Gulumbe, “Identifying the limitation of stepwise selection for variable selection in regression analysis,” *American Journal of Theoretical and Applied Statistics*, vol. 4, no. 5, pp. 414–419, 2015.
- [88] R. Routledge, “When stepwise regression fails: Correlated variables some of which are redundant,” *International Journal of Mathematical Education in Science and Technology*, vol. 21, no. 3, pp. 403–410, 1990.
- [89] M. S. Lewis-Beck, “Stepwise regression: A caution,” *Political Methodology*, pp. 213–240, 1978.
- [90] T. R. Gault *et al.*, “Comparing performance of standard and iterative linear unmixing methods for hyperspectral signatures,” in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII*, SPIE, vol. 9840, 2016, pp. 579–590.
- [91] R. B. Mhenni *et al.*, “Spectral unmixing with sparsity and structuring constraints,” in *2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, IEEE, 2018, pp. 1–5.
- [92] C. A. Bateson, G. P. Asner, and C. A. Wessman, “Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis,” *IEEE transactions on geoscience and remote sensing*, vol. 38, no. 2, pp. 1083–1094, 2000.
- [93] A. Marinoni and P. Gamba, “A novel approach for efficient p -linear hyperspectral unmixing,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1156–1168, 2015.
- [94] S. Jokar and M. E. Pfetsch, “Exact and approximate sparse solutions of underdetermined linear equations,” *SIAM Journal on Scientific Computing*, vol. 31, no. 1, pp. 23–44, 2008.
- [95] F. Schmidt *et al.*, “Spectral unmixing with positivity, sparsity and group exclusivity constraints,” Copernicus Meetings, Tech. Rep., 2021.

- [96] F. Schmidt *et al.*, “Mineralogy analysis using linear unmixing under group constraint,” in *In 2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, IEEE, 2023.
- [97] A. Dache *et al.*, “Exact and heuristic methods for simultaneous sparse coding,” in *2023 31st European Signal Processing Conference (EUSIPCO)*, IEEE, 2023, pp. 1753–1757.
- [98] S. Ermon *et al.*, “Pattern decomposition with complex combinatorial constraints: Application to materials discovery,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 29, 2015.
- [99] S. Bourguignon *et al.*, “Exact sparse approximation problems via mixed-integer programming: Formulations and computational performance,” *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1405–1419, 2015.
- [100] D. Madigan and A. E. Raftery, “Model selection and accounting for model uncertainty in graphical models using occam’s window,” *Journal of the American Statistical Association*, vol. 89, no. 428, pp. 1535–1546, 1994.
- [101] J. Preston and W. Basener, “Theoretical and practical progress in hyperspectral pixel unmixing with large spectral libraries from a sparse perspective,” *arXiv preprint arXiv:2408.07580*, 2024.
- [102] J. Preston and W. Basener, “Hyperspectral unmixing using iterative, sparse and ensembling approaches for large spectral libraries applied to soils and minerals,” *arXiv preprint arXiv:2503.16298*, 2025.
- [103] J. Preston and W. Basener, “Spectral unmixing comparison with sparse, iterative and mixed integer programming models,” *arXiv preprint arXiv:2503.17118*, 2025.
- [104] A. Gramfort, *Scikit learn linear regression*, https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html, [Online; accessed 7-July-2024], Jul. 2024.

- [105] *Scipy optimize nnls*, <https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.nnls.html>, [Online; accessed 7-July-2024].
- [106] *Ridge*, https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Ridge.html, [Online; accessed 7-July-2024], Jul. 2024.
- [107] *Lasso*, https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html, [Online; accessed 7-July-2024], Jul. 2024.
- [108] B. Basener, “An automated method for identification and ranking of hyperspectral target detections,” in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*, SPIE, vol. 8048, 2011, pp. 131–138.

A Appendix

Category	Technique	RMSE Mean	Model Size	Runtime (s)	Detection	Mean Precision
Non-sparse Approaches	OLS	0.0014	481	0.0017	1.0	0.0351
	Ridge	0.0980	228	0.0040	1.0	0.1322
	NNLS	0.0036	13	0.0011	1.0	0.2756
Sparse Approaches	LASSO	0.0226	8	0.1261	1.0	0.4145
	ElasticNet	0.0233	13	0.0973	1.0	0.4388
Optimization Approaches	DFS NNLS	0.0080	4	2.3625	1.0	0.5826
	DFS LASSO	0.0391	4	2.2703	1.0	0.5495
	DFS ElasticNet	0.0386	5	2.2638	1.0	0.5161
	BFS NNLS	0.0192	4	1.554	0.31	0.1308
	MINLP	0.0044	4	134.3793	1.0	0.3599
Ensembling Approaches	BMA NNLS	0.0102	2	0.0779	1.0	0.7651
	BMA LASSO	0.0119	2	0.6488	0.95	0.6202
	BMA ElasticNet	0.0117	2	0.7823	0.96	0.6015
Nonlinear Approaches	BMAQ NNLS	0.0744	3	0.0793	1.0	0.7141
	BMAQ LASSO	0.0667	3	0.8632	0.95	0.6077
	BMAQ ElasticNet	0.0559	2	0.8797	0.95	0.6166
Band Decorrelation	HySUDeB	0.1984	6	1.9236	0.95	0.6032

(a) Metrics for unmixing techniques for the alunite ROI.

Category	Technique	RMSE Mean	Model Size	Runtime (s)	Detection	Mean Precision
Non-sparse Approaches	OLS	0.0012	481	0.0030	1.0	0.0424
	Ridge	0.1210	351	0.0057	1.0	0.1109
	NNLS	0.0044	11	0.0012	1.0	0.3656
Sparse Approaches	LASSO	0.0311	9	0.0311	1.0	0.3909
	ElasticNet	0.0439	15	0.1800	1.0	0.4727
Optimization Approaches	DFS NNLS	0.0080	4	2.2521	0.65	0.3089
	DFS LASSO	0.0200	4	2.3054	0.66	0.2847
	DFS ElasticNet	0.0203	5	2.4664	0.67	0.2490
	BFS NNLS	0.0144	5	3.8240	0.49	0.1226
	MINLP	0.0047	2	94.7523	1.0	0.5102
Ensembling Approaches	BMA NNLS	0.0116	3	0.0636	0.42	0.2601
	BMA LASSO	0.0136	3	0.9077	0.85	0.5159
	BMA ElasticNet	0.0133	3	0.8784	0.85	0.5232
Nonlinear Approaches	BMAQ NNLS	0.1339	3	0.1887	0.59	0.2660
	BMAQ LASSO	0.1588	3	1.0417	0.49	0.2697
	BMAQ ElasticNet	0.1178	3	1.0427	0.74	0.2919
Band Decorrelation	HySUDeB	0.2131	6	2.5341	0.94	0.4623

(b) Metrics for unmixing techniques for the kaolinite ROI.

Category	Technique	RMSE Mean	Model Size	Runtime (s)	Detection	Mean Precision
Non-sparse Approaches	OLS	0.0012	481	0.0037	1.0	0.0219
	Ridge	0.0582	380	0.0044	1.0	0.0511
	NNLS	0.0031	14	0.0014	0.72	0.0633
Sparse Approaches	LASSO	0.0604	7	0.0731	0.85	0.1394
	ElasticNet	0.0345	11	0.0915	1.0	0.1690
Optimization Approaches	DFS NNLS	0.0049	4	1.9409	0.56	0.1415
	DFS LASSO	0.0281	4	1.9994	0.56	0.1507
	DFS ElasticNet	0.0241	4	1.9951	0.57	0.1593
	BFS NNLS	0.0091	6	2.6035	0.27	0.0359
	MINLP	0.0037	4	143.7650	0.77	0.1815
Ensembling Approaches	BMA NNLS	0.0091	3	0.1136	0.21	0.0979
	BMA LASSO	0.0130	3	0.9359	0.52	0.3180
	BMA ElasticNet	0.0129	3	0.9344	0.52	0.2867
Nonlinear Approaches	BMAQ NNLS	0.1679	3	0.1081	0.0	0.0
	BMAQ LASSO	0.1781	3	0.8320	0.27	0.1847
	BMAQ ElasticNet	0.1602	3	0.6974	0.30	0.1734
Band Decorrelation	HySUDeB	0.2089	7	2.5323	0.22	0.0347

(c) Metrics for unmixing techniques for the Montmorillonite ROI.

Table A.1: Complete comparison of unmixing techniques for each of the three ROIs.