

# Spectral Umixing Comparison with Sparse, Iterative and Mixed Integer Programming Models

Jade F. Preston<sup>1†</sup> and William F. Basener<sup>1\*†</sup>

<sup>1</sup>School of Data Science, University of Virginia, Charlottesville, VA.

\*Address correspondence to: wb8by@virginia.edu

## Abstract

Hyperspectral unmixing is the analytical process of determining the pure materials and estimating the proportions of such materials composed within an observed mixed pixel spectrum. We can unmix mixed pixel spectra using linear and nonlinear mixture models. Ordinary least squares (OLS) regression serves as the foundation for many linear mixture models employed in Hyperspectral Image analysis. Though variations of OLS are implemented, studies rarely address the underlying assumptions that affect results. This paper provides an in depth discussion on the assumptions inherently endorsed by the application of OLS regression. We also 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.

## 1 Introduction

Hyperspectral images capture the electromagnetic radiation reflected from surfaces and record the reflectance across hundreds or thousands of contiguous spectral bands [1]. Researchers refer to this mapping of material reflectance across spectral bands as spectral signatures or spectra. Pixels within the images frequently contain a mixture of different materials – ultimately modifying the amount of light photons returning to the hyperspectral sensor [2]. Researchers aim to identify and quantify these materials through a process called spectral unmixing. Material mixtures can be categorized and modeled as linear or nonlinear depending on compositional complexity [3].

The linear mixture model (LMM) – depicted by equation 1 – is often interpreted as a regression problem [3, 4]. The observed mixed pixel spectrum (singular form of spectra) is represented by the dependent variable,  $\mathbf{y}$ , and can be further described as a vector of the pixel’s recorded reflectance at waveband ( $i$  to  $N$ ). We, therefore, model the observed spectrum by inferring the pure materials

or endmembers from an image or constituent spectra from a spectral library. In this study, we use an extensive spectral library for unmixing rather relying on “pure” image pixels. The matrix of constituent spectra is represented by  $\mathbf{S}$ . Neither the image endmembers nor the abundances are known *a priori*. However, if the spectral library is small, we can invert the library to estimate these unknown abundances, represented by the vector  $\mathbf{a}$  [5]. Inversion becomes impossible when the spectral library is large — an increasingly common scenario — particularly when there are more library spectra than spectral bands. Equation 2 displays the unmixing solution estimating the unknown, fractional proportion of materials. All models inherently incorporate some level of error. Thus, the unmixing solution is an estimate indicated by the lack of error vector,  $\mathbf{E}$ , shown in equation 1. In least squares models, the error is referred to as the residual or in this case the distance between the inferred reflectance and the true pixel reflectance at wavelength  $i$ .

$$\mathbf{y} = \mathbf{Sa} + \mathbf{E} \quad (1)$$

$$\hat{\mathbf{a}} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{y} \quad (2)$$

The basic form of linear regression is also referred to as OLS [6, 7]. OLS minimizes modeling error by reducing the sum of the squared differences between the observed and inferred reflectance values. There are alternative regression approaches (i.e. generalized or weighted least squares), but this study focuses on OLS because it serves as the baseline for most unmixing techniques [5]. A common metric for assessing model error is mean squared error (MSE), which is calculated by normalizing the residual sum of squares by the number of data points. Equation 3 presents the formula for MSE. The OLS objective is to minimize the MSE by identifying the set of coefficients that maximize the similarity between the response and the regressor values [8].

$$\text{MSE} = \frac{1}{M} \sum_{i=1}^M (\mathbf{y}_i - (\mathbf{Sa})_i)^2 \quad (3)$$

The scope of this study is primary material identification via four unmixing methods. We aim to identify specific substances embedded in a mixed pixel regardless of pixel composition. In this paper, we will examine the assumptions underlying OLS and their alignment, or lack thereof, with the LMM. For instance, the OLS mixture model fails to provide practical abundance estimates when the spectral library is large.

This paper also provides a comparison focusing on four highly effective unmixing techniques: sparse regression, depth-first search (DFS) feature selection, mixed integer nonlinear programming (MINLP), and band decorrelation. In many past papers, the implementation of certain feature selection techniques — i.e as step-wise regression — have been critiqued. It is our additional aim to discuss the attributes of iterative feature selection that are useful in spectral unmixing.

Dimension reduction and decorrelation of library spectra are common practice in spectral unmixing. To our knowledge, no other study has pre-processed the data by whitening the spectral bands as part the unmixing methodology. This paper presents a novel approach called Hyperspectral unmixing with decorrelated bands (HySUDeB). Each of the techniques discussed in this paper

respond to one or more of the inadequacies realized through OLS unmixing. We compare the three previously mentioned unmixing techniques to HySUDeB and investigate their performance trade offs. This investigation includes a description of the physical-chemical structure of the materials aiding in accurate detection.

## 1.1 Ordinary Least Square Assumptions

Researchers, in the spectral analysis field, often discuss the assumptions supporting the LMM but the assumptions behind OLS are rarely explained. The outline below highlights primary OLS assumptions and discusses their validity in the context of spectral unmixing [5, 9, 10]:

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) in which both the contribution of the input variables and the error term are considered additive.

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 high vegetation or mineral deposits) [3, 11].

2. Independent Identically Distributed (IID) error — The residuals from OLS are assumed to satisfy three key properties: independence, homoscedasticity, and normality with mean centered at zero. Independence implies that there is no correlation between the residuals. This independence assumption means that 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). Homoscedasticity means the variance or spread of the residuals is constant across the input variables. When we say the residuals are normally distributed with mean equaling zero, it means the residuals follow a bell-shaped curve when plotted as a histogram, and their expected value is zero. Because the expected value of the error is zero, this implies: 1. OLS is an unbiased estimator and 2. there is no correlation between the residuals and the input variables. Therefore, we assume error is not computed in any one direction.

In the context of the LMM, a multivariate normal distribution is assumed for the residuals with the mean centered at the observed pixel spectrum. 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 [12, 13]. Factors such as correlated bands, spectral variability, and spatio-temporal effects can disrupt homoscedasticity. As a

result, residuals may exhibit fluctuating variances or in certain cases follow entirely different distributions [14].

3. Noncollinear Regressor Variables — OLS requires the regressor variables exhibit noncollinearity, meaning there is no linear dependency between the 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 2.

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

4. Correctly established models — The input variables are considered fixed implying that the information available to describe the response variable is 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.

## 1.2 Practical LMM Constraints

The unconstrained OLS model provides unrealistic abundance results and requires the implementation of one or more practical constraints. The two primary constraints enforced are the nonnegative abundance and abundance sum to one constraints [15, 16]. Negative abundance observed from the unconstrained model stem from a combination of over-fitting of the observed variable and the minimal presences of a particular substance [17]. The nonnegative abundance constraint – commonly referred to as nonnegative least squares (NNLS) – ensures any variables included in the model have coefficients greater than or equal to zero. Equation 4 displays the nonnegative abundance constraint regularly imposed on the OLS model.  $N$  represents the explanatory variables. In the paper by Keshava et al., the authors discuss the implementation of the sum-to-one abundance constraint. This constraint forces the estimated coefficients of the explanatory variables to sum to one, adhering to the practical assertion that the materials in the model make up a fractional proportion of the area in the pixel [3]. Within this paper, we relax this constraint to comply with an alternative assumption: materials in a mixed pixel may appear brighter in certain scenarios causing the abundances to

sum greater than one. Thus, the only practical constraint we enforce in this study is nonnegative abundances.

$$a_i \geq 0 \quad \forall 1 \leq i \leq N \quad (4)$$

### 1.3 Alternative Unmixing Techniques

#### 1.3.1 Sparse Regression

Sparse unmixing is a technique in spectral unmixing that addresses some of the misalignments found with OLS assumptions. High correlation between library spectra is a significant challenge in spectral unmixing. Large spectral libraries are essential to capture spectral variability in intimate mixtures, but their use can lead to overfitting and an ill-posed inversion problem, thereby making regularization necessary. Sparse unmixing addresses these issues by selecting the optimal subset of input variables from the library to describe the observed spectrum [4]. This approach introduces regularization to the OLS model through penalties on the parameters.

There are many studies that have achieved unmixing success using sparse unmixing approaches. The common penalties imposed are the  $L_1$  norm (also known as LASSO regression),  $L_2$  norm (also known as RIDGE regression) or a combination of the  $L_1$  and  $L_2$  norm (their combination is known as ElasticNet). LASSO regression uses Manhattan distancing to shrink coefficients to zero penalizing the size of the model. Subsequently, this  $L_1$  penalty adheres to the nonnegative abundance constraint. Equation 5 displays the modified version of equation 3 which includes the  $L_1$  penalty on the model coefficients. The set of coefficients,  $\hat{\mathbf{a}}$ , minimizes the objective function subject to the  $L_1$  penalty norm.

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \left( \frac{1}{M} \|\mathbf{y} - \mathbf{S}\mathbf{a}\|_2^2 + \lambda \|\mathbf{a}\|_1 \right) \quad (5)$$

#### 1.3.2 Iterative Approaches

An alternative to sparse approaches are iterative approaches i.e. depth- and breadth-first search strategies. Depth- and breadth-first search strategies are often incorporated in graph algorithms or decision trees [18, 19], but their underlying principles have been adapted to a wide variety of use cases i.e. step-wise regression. Both strategies iteratively identify a subset of features that best fit the data [20]. Step-wise regression is primarily implemented for model selection with a secondary aim of parameter estimation. In the context of unmixing, step-wise regression responds to the goal of sparse unmixing which is to identify the subset of explanatory variables from a far larger set of materials [21, 22].

Many statisticians caution against the implementation of step-wise regression due to issues with multi-collinearity. When explanatory variables are highly correlated, step-wise 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 [23]. 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 [24].

The studies discussing the limitations of step-wise regression often exhibit two characteristics that are less relevant to spectral unmixing. Many of the correlated variables in spectral libraries are the variations of the same substance rather than two highly correlated but different variables. Additionally, OLS variables can have a negative impact on the response variable. Practicality prohibits the inference of negative abundance coefficients in the LMM. Considering these differences, stepwise regression's disadvantages become less pronounced in the spectral unmixing context. The primary aim of variable selection aligns with stepwise regression's design, making it well-suited to identifying the most statistically significant materials influencing pixel reflectance. Winter et al. explains that the iterative process of step-wise regression coupled with nonnegative abundance enforcement lowers modeling error [17]. In another paper, Gault et al. compares NNLS, sparse and step-wise approaches. The authors explain that step-wise regression can out perform non-iterative approaches especially when detection is the main goal [25].

With these considerations in mind, we implement an iterative search strategy, focusing specifically on DFS feature selection. This approach begins with an empty set, iteratively identifies the most significant explanatory variables and incorporates 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 then used to calculate the probability (p-value) of obtaining the observed F-statistic by chance. A low p-value suggests that the observed 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 [22–24].

### 1.3.3 MINLP

Despite its potential for higher accuracy [26], 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 [27, 28], integer programming [29, 30], mixed integer programming (MIP), MINLP [31, 32] and other variations of nonlinear programming [33–35]. 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 objective – a minimization or maximization function – composed of decision variables. These variables are iteratively adjusted within the defined limits to determine the best solution. The limits imposed are called constraints and represent practical requirements or boundaries. MIP integrates concepts from both linear and integer programming because the decision variables are integer and continuous. MINLP is an optimization method originating from MIP, but incorporates nonlinear relationships [36].

Bateson et al. completed a study in which they developed endmember bundles using principal component analysis and simulated annealing to determine the dimension and vertices of the data simplex. Each bundle is a collection of spectra for a single material; the author incorporated grass, tree and soil bundles. Once the bundles were created, Bateson et al. unmixed pixels using linear programming [27]. In another study, Bourguignon et al. compare nine variations of sparse mixture models, showing the equivalence of more traditional mixture models to MIPs. The authors concluded that with an accepted increase in computational demand, mixed integer formulations can reduce the error of abundance estimation and provide exact solutions [35]. Mhenni et al. echo these sentiments in their paper where they conduct unmixing using a mathematical program incorporating sparse and exclusive group constraints [26].

## 2 Materials and Methods

### 2.1 Experimental Design

Our observed dataset was derived from a hyperspectral image captured using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). AVIRIS is one of the first airborne hyperspectral imagers developed in the 1980s by NASA’s Jet Propulsion Laboratory [37]. It currently collects data across 224 contiguous spectral bands, covering wavelengths from 0.38 to 2.5 micrometers. The image used in this study consists of 350x400 pixels, capturing a section of Cuprite Hills, Nevada [38]. Cuprite Hills is a well-known mineralogical site and a prime testing ground for studying natural intimate mixtures. Additionally, we utilized the spectral library from the U.S. Geological Survey (USGS) to infer materials embedded in the observed Cuprite image pixels. Our spectral library was comprised of 481 mineral spectra. Many of the spectra from the library are a variation of the same substance to aid material detection within the intimate mixtures found at Cuprite Hills [39].

The specific regions of interest (ROIs) within the AVIRIS image are comprised of pixels from alunite, kaolinite and montmorillonite locations. We used a graphical user interface, “Hyperspectralpy” [40] 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 applies a whitening transformation to the image and library spectra. Following the whitening process, an ACE score is computed depicting the strength of the target mineral’s presence within image. Figure 1 displays the selected ROIs via the red colored regions. The top row in Figure 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 1).

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.

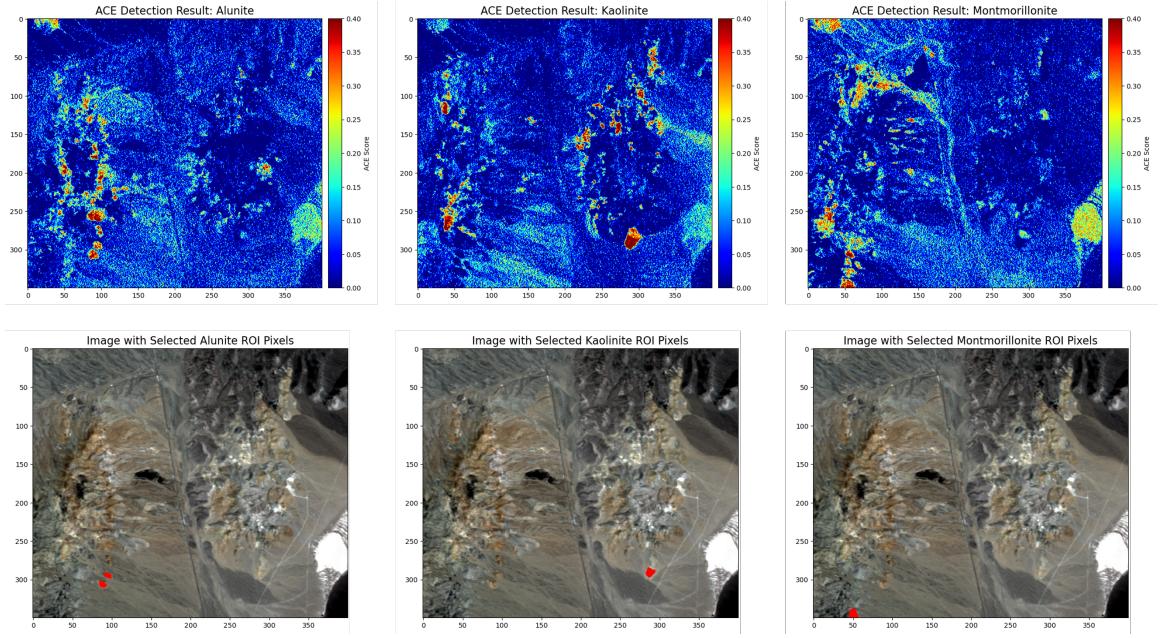


Figure 1: Display of the selected ROIs for three target minerals: alunite, kaolinite, and montmorillonite. The top row of plots show the regions where the target mineral shows higher abundance and the bottom row depicts the actual selected pixels from the locations with higher presence of the target mineral.

Figure 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 a  $\text{SO}_4^{2-}$  chemical formula because they are comprised of one sulfur and four oxygen atoms. Tectosilicates have a  $\text{SiO}$  chemical formula which is indicative of their silicon and oxygen composition. Nesosilicates generally incorporate a silicon atom and four oxygen atoms,  $\text{SiO}_4$ . Phyllosilicates have a chemical formula,  $\text{Si}_2\text{O}_5$ , describing their composition of two silicon atoms and five oxygen atoms.

## 2.2 Statistical Analysis

### 2.2.1 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.

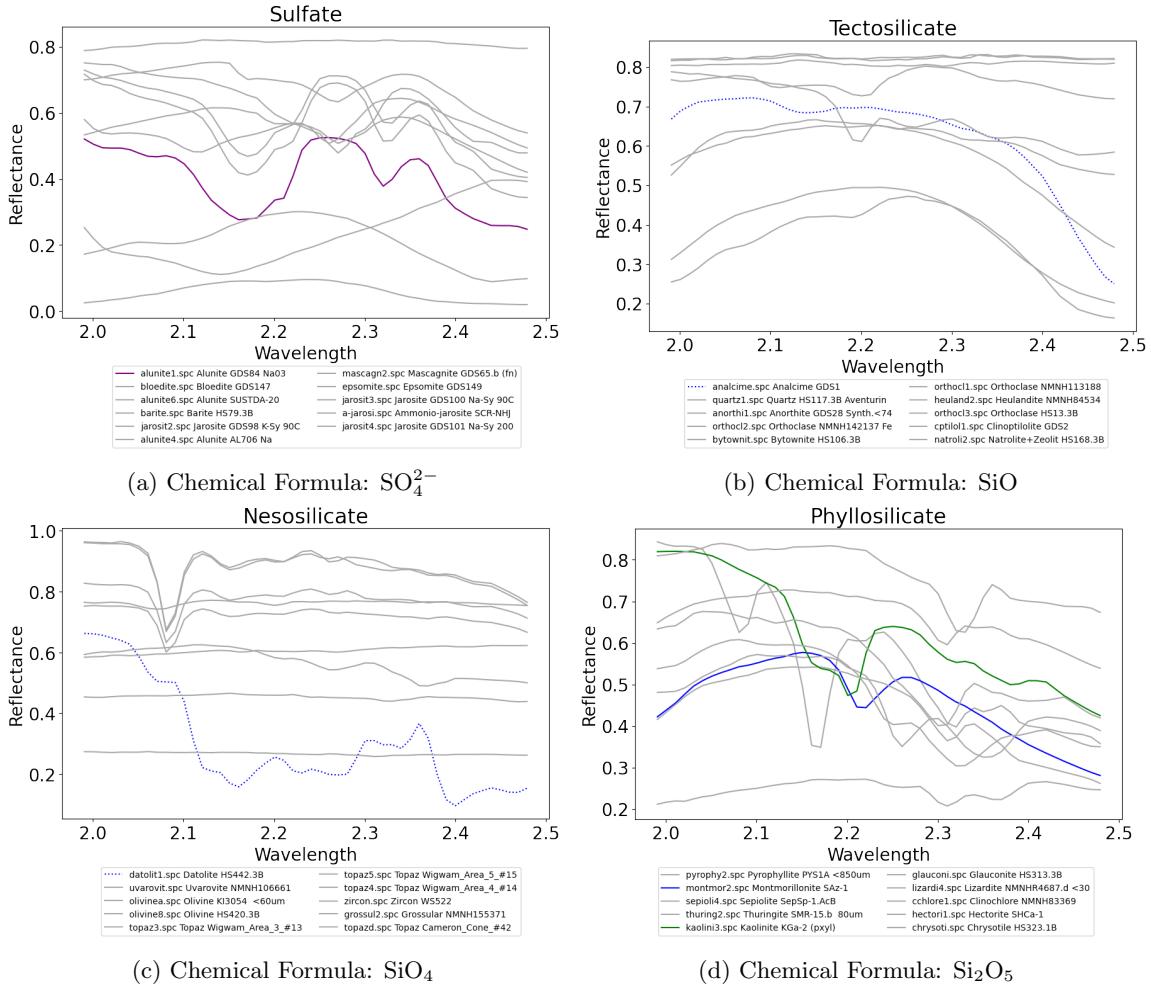


Figure 2: Comparison of the mineral types sulfate, tectosilicate, and nesosilicate, and phyllosilicate. The purple spectrum (Figure 2a) represents alunite. The green and blue spectra (Figure 2d) represent kaolinite and montmorillonite, respectively. The dotted blue spectra within the other plots (Figures 2b and 2c) are minerals that were often incorporated in the models.

### 2.2.2 Depth-first Search 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.

### 2.2.3 Mixed Integer Nonlinear Program

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 [26, 35]. 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

$N$  = number of library spectra

$B_i$  = 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 ( $s_i$ ) is binary. The abundance associated with a spectrum ( $a_i$ ) must be positive.

### 2.2.4 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 originally 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 [41]. Once band decorrelation was complete, we applied LASSO regression to the whitened spectra.

## 3 Results

We compared the unmixing performance of each technique based on computation time, error, detection accuracy and mean average precision at  $k$ . 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. Figure 3 displays each technique's performance for the alunite ROI in terms of computation time and root mean squared error (RMSE) – square root of equation 3. Though Figure 3 is the result for the alunite ROI and the resulting plots for kaolinte and montmorillonite have different numeric values, the performance spread 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.

The results for detection accuracy are shown in Table 1. For each of the ROIs, LASSO regression incorporated the 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,

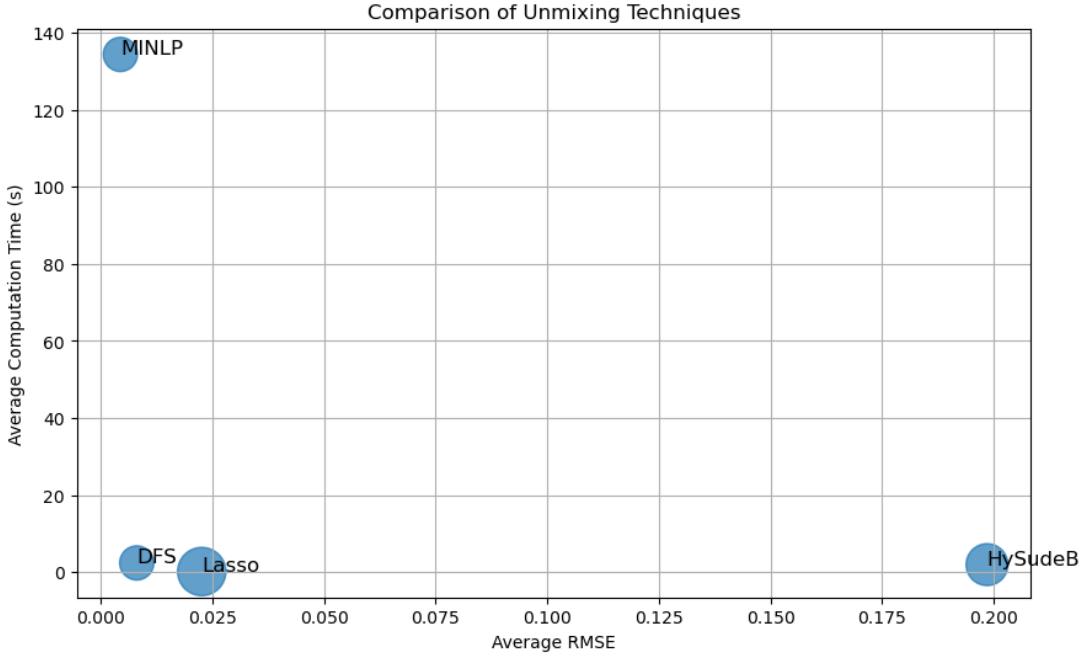


Figure 3: Display of technique performance based on average error and computation time for the alunite ROI.

MINLP only included the target mineral once – while other approaches incorporated variations of the target in many models.

Table 2 displays the results in terms of detection precision. Precision at  $k$  is an indication 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 preformed the best with the other two regions, kaolinite and montmorillonite.

	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 1: Detection percentages for target 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 the each modeling approach. The most commonly included categories regardless of technique were sulfates, tectosilicates, nesosilicates, carbonates and phyllosilicates. Figure 4 shows the count for all the included mineral categories across all three ROIs. The categories shown in Figure 4 are

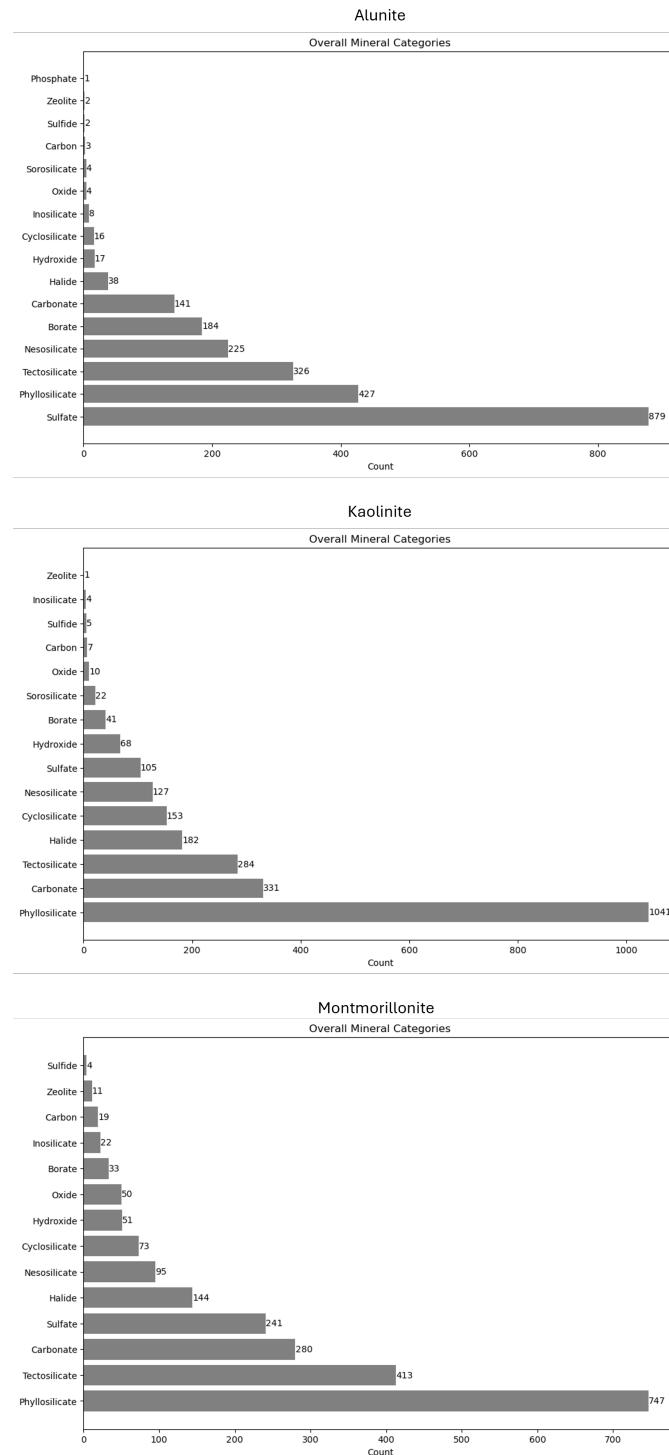


Figure 4: Display of the count of the included minerals across all the implemented approaches.

Mean Average Precision at k

	Lasso	DFS	MINLP	HySudeB
Alunite	0.41	0.58	0.36	0.60
Kaolinite	0.39	0.31	0.54	0.46
Montmorillonite	0.14	0.14	0.18	0.03

Table 2: Average mean precision at k for minerals (alunite, kaolinite, and montmorillonite) across four techniques: Lasso, DFS, MINLP, and HySudeB.

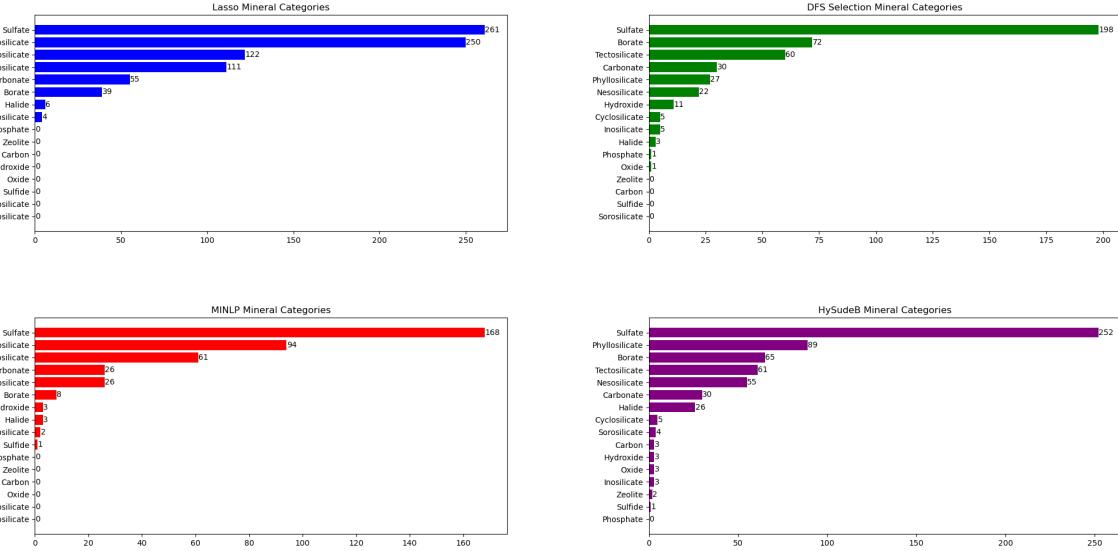


Figure 5: Display of the count of the minerals from the library that were included in the models for the alunite ROI.

not an exhaustive list of the mineral types found in the spectral library. Figure 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 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.

## 4 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 are 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 in every category. 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. LASSO regression performed the best due to cross validation. The penalty parameter allowed the model to become less sparse. This lowered sparsity ensured: 1) the target mineral was included in the model at a minimum 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 are 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.

## Acknowledgments

The authors would like to thank the School of Data Science at the University of Virginia for their support during this research process and providing access to publication catalogs.

## References

1. Manolakis D, Marden D, Shaw GA, et al. Hyperspectral image processing for automatic target detection applications. *Lincoln laboratory journal* 2003;14:79–116.
2. Somers B, Asner GP, Tits L, and Coppin P. Endmember variability in spectral mixture analysis: A review. *Remote Sensing of Environment* 2011;115:1603–16.
3. Keshava N and Mustard JF. Spectral unmixing. *IEEE signal processing magazine* 2002;19:44–57.

4. Borsoi RA, Imbiriba T, Bermudez JCM, et al. Spectral variability in hyperspectral data unmixing: A comprehensive review. *IEEE geoscience and remote sensing magazine* 2021;9:223–70.
5. Dismuke C and Lindrooth R. Ordinary least squares. Methods and designs for outcomes research 2006;93:93–104.
6. Fox J. Applied regression analysis and generalized linear models. Sage publications, 2015.
7. Burton AL. OLS (Linear) regression. *The encyclopedia of research methods in criminology and Criminal Justice* 2021;2:509–14.
8. Wang Z and Bovik AC. Mean squared error: Love it or leave it? A new look at Signal Fidelity Measures. *IEEE Signal Processing Magazine* 2009;26:98–117.
9. Poole MA and O'Farrell PN. The assumptions of the linear regression model. *Transactions of the Institute of British Geographers* 1971:145–58.
10. Tranmer M and Elliot M. Multiple linear regression. *The Cathie Marsh Centre for Census and Survey Research (CCSR)* 2008;5:1–5.
11. Dobigeon N, Tourneret JY, Richard C, Bermudez JCM, McLaughlin S, and Hero AO. Nonlinear unmixing of hyperspectral images: Models and algorithms. *IEEE Signal processing magazine* 2013;31:82–94.
12. Chang CI. Hyperspectral data processing: algorithm design and analysis. John Wiley & Sons, 2013.
13. Chang CI and Liu KH. Progressive band selection of spectral unmixing for hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing* 2013;52:2002–17.
14. Bioucas-Dias JM, Plaza A, Camps-Valls G, Scheunders P, Nasrabadi N, and Chanussot J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing magazine* 2013;1:6–36.
15. Keshava N, Kerekes JP, Manolakis DG, and Shaw GA. Algorithm taxonomy for hyperspectral unmixing. In: *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*. Vol. 4049. SPIE. 2000:42–63.
16. Heinz DC et al. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE transactions on geoscience and remote sensing* 2001;39:529–45.
17. Winter ME, Lucey PG, and Steutel D. Examining hyperspectral unmixing error reduction due to stepwise unmixing. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX*. Vol. 5093. SPIE. 2003:380–9.
18. Kozen DC and Kozen DC. Depth-first and breadth-first search. The design and analysis of algorithms 1992:19–24.
19. Bundy A and Wallen L. Breadth-first search. Catalogue of artificial intelligence tools 1984:13–3.
20. Dash M and Liu H. Feature selection for classification. *Intelligent data analysis* 1997;1:131–56.

21. Ruengvirayudh P and Brooks GP. Comparing stepwise regression models to the best-subsets models, or, the art of stepwise. *General linear model journal* 2016.
22. Olusegun AM, Dikko HG, and Gulumbe SU. Identifying the limitation of stepwise selection for variable selection in regression analysis. *American Journal of Theoretical and Applied Statistics* 2015;4:414–9.
23. Routledge R. When stepwise regression fails: correlated variables some of which are redundant. *International Journal of Mathematical Education in Science and Technology* 1990;21:403–10.
24. Lewis-Beck MS. Stepwise regression: A caution. *Political Methodology* 1978;2:13–40.
25. Gault TR, Jansen ME, DeCoster ME, Jansing ED, and Rodriguez BM. Comparing performance of standard and iterative linear unmixing methods for hyperspectral signatures. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII*. Vol. 9840. SPIE. 2016:579–90.
26. Mhenni RB, Bourguignon S, Ninin J, and Schmidt F. Spectral unmixing with sparsity and structuring constraints. In: *2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE. 2018:1–5.
27. Bateson CA, Asner GP, and Wessman CA. Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis. *IEEE transactions on geoscience and remote sensing* 2000;38:1083–94.
28. Marinoni A and Gamba P. A novel approach for efficient  $p$ -linear hyperspectral unmixing. *IEEE Journal of Selected Topics in Signal Processing* 2015;9:1156–68.
29. Jokar S and Pfetsch ME. Exact and Approximate Sparse Solutions of Underdetermined Linear Equations. *SIAM Journal on Scientific Computing* 2008;31:23–44.
30. Liu Y, Condessa F, Bioucas-Dias JM, Li J, Du P, and Plaza A. Convex formulation for multi-band image classification with superpixel-based spatial regularization. *IEEE Transactions on Geoscience and Remote Sensing* 2018;56:2704–21.
31. Schmidt F, Salomon G, Mège D, Gurgurewicz J, and Bourguignon S. Spectral unmixing with positivity, sparsity and group exclusivity constraints. Tech. rep. Copernicus Meetings, 2021.
32. Schmidt F, Salomon G, Bourguignon S, Gurgurewicz J, and Mège D. 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.
33. Dache A, Vandaele A, Gillis N, and Nadisic N. Exact and Heuristic Methods for Simultaneous Sparse Coding. In: *2023 31st European Signal Processing Conference (EUSIPCO)*. IEEE. 2023:1753–7.
34. Ermon S, Le Bras R, Suram S, et al. Pattern decomposition with complex combinatorial constraints: Application to materials discovery. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 29. 1. 2015.

35. Bourguignon S, Ninin J, Carfantan H, and Mongeau M. Exact sparse approximation problems via mixed-integer programming: Formulations and computational performance. *IEEE Transactions on Signal Processing* 2015;64:1405–19.
36. Smith JC and Taskin ZC. A tutorial guide to mixed-integer programming models and solution techniques. *Optimization in medicine and biology* 2008;521–48.
37. Qian SE. Hyperspectral satellites, evolution, and development history. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2021;14:7032–56.
38. Swayze GA, Clark RN, Goetz AF, et al. Mapping advanced argillic alteration at Cuprite, Nevada, using imaging spectroscopy. *Economic geology* 2014;109:1179–221.
39. Clark R, Swayze G, Wise R, 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.
40. Basener B. Hyperspectralpy. 2022. URL: <https://pypi.org/project/hyperspectralpy/>.
41. Basener B. An automated method for identification and ranking of hyperspectral target detections. In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*. Vol. 8048. SPIE. 2011:131–8.

## A Referenced code

All the referenced and supporting code will be stored within the corresponding GitHub repository (<https://github.com/bakerjf1993/Hyperspectral-Unmixing/tree/main/2024%20Journal%20of%20Remote%20Sensing>).