

Article

A Robust Stochastic Approach to Mineral Hyperspectral Analysis for Geometallurgy

Álvaro F. Egaña [†], Felipe A. Santibáñez-Leal [†], Christian Vidal, Gonzalo Díaz, Sergio Liberman and Alejandro Ehrenfeld ^{*,†}

Advanced Laboratory for Geostatistical Supercomputing (ALGES), Advanced Mining Technology Center (AMTC), Department of Mining Engineering, University of Chile, Santiago 8370451, Chile; aegana@alges.cl (A.F.E.); fsantibanez@alges.cl (F.A.S.-L.); cvidal@alges.cl (C.V.); gdiaz@alges.cl (G.D.); sliberman@alges.cl (S.L.)

* Correspondence: aehrenfeld@alges.cl

† These authors contributed equally to this work.

Received: 31 October 2020; Accepted: 11 December 2020; Published: 18 December 2020



Abstract: Most mining companies have registered important amounts of drill core composite spectra using different acquisition equipment and by following diverse protocols. These companies have used classic spectrography based on the detection of absorption features to perform semi-quantitative mineralogy. This methodology requires ideal laboratory conditions in order to obtain normalized spectra to compare. However, the inherent variability of spectral features—due to environmental conditions and geological context, among others—is unavoidable and needs to be managed. This work presents a novel methodology for geometallurgical sample characterization consisting of a heterogeneous, multi-pixel processing pipeline which addresses the effects of ambient conditions and geological context variability to estimate critical geological and geometallurgical variables. It relies on the assumptions that the acquisition of hyperspectral images is an inherently stochastic process and that ore sample information is deployed in the whole spectrum. The proposed framework is basically composed of: (a) a new hyperspectral image segmentation algorithm, (b) a preserving-information dimensionality reduction scheme and (c) a stochastic hierarchical regression model. A set of experiments considering white reference spectral characterization and geometallurgical variable estimation is presented to show promising results for the proposed approach.

Keywords: geometallurgical characterization; hyperspectral processing and analysis; hyperspectral clustering; stochastic modeling

1. Introduction

Mining industry has taken advantage of near infrared (NIR) spectrography [1] technology to characterize geological and metallurgical samples. They record drill core trays using hyperspectral single pixel scanners, such as Hylogger [2], or 2D imaging systems such as Corescan [3,4], and they acquire spectra from composites of drill core segments using single pixel hyperspectral instruments such as ASD SpecLab [5–7]. In the first case, single pixel spectral analysis is used for semi-quantitative mineralogy, lithology and alteration characterization, supported by several mature unmixing techniques or linear regressions [8,9]. In the second case, spectra are used mainly for clays and certain mineral identification tasks, based on classical NIR analysis [1,10].

This large amount of available information has produced relevant challenges and difficulties, which gave us our initial motivation to work firstly with machine learning approaches, with some success for geological sample analysis applications, and more recently with a novel stochastic approach for metallurgical sample analysis applications.

From the spectral characterization point of view, it is possible to organize mineral samples into two categories. The difference between them resides in their contents. A specific sample will fall into one category or the other depending on whether it is composed of mainly rocks or comminuted minerals, both with different grain profiles within their own scales. For the sake of clarity, in this paper we will call these categories geological and metallurgical samples respectively. These definitions are relevant because they represent different challenges for hyperspectral analysis.

In geological samples there are surface textural features that give information about rock formation processes or geological variables such as lithology, alteration or styles of mineralization. Moreover, a rock sample surface provides valuable information about its origin and history—we will deepen this concept in Section 2.4. The study of the relationship and disposition of its components allows us to understand (a) the surroundings where it was deposited, (b) what the context during the rock formation process was and (c) interactions that took place among its components. On the other hand, in metallurgical samples, mineral particles coming from different parts of the original rocks are mixed. Spectral variability among pixels becomes very relevant in this case.

Section 2 presents background about variability that usually arises in a hyperspectral image sample spectrum. Aspects of our work relevant to texture analysis in the geological samples, including spatial context, color and spectral context analysis, are described as well. Then, background on dimensionality reduction and hyperspectral image clustering methods is presented.

A stochastic approach to hyperspectral analysis of metallurgical samples is proposed in Section 4, where a general methodology and processing pipelines that consider hyperspectral data as a proxy for the estimation of geometallurgical variables are provided. Main pipeline tools are described in Section 5, considering dimensional reduction, clustering and stochastic hierarchical regression.

Demonstrative experiments and results are shown in Section 6, and a final discussion is presented in Section 7.

2. Background

2.1. Laboratory and Field Based Geological Hyperspectral Analysis

A hyperspectral image (HSI) could be considered as a set of high-dimensional signals containing massive amounts of information, because of both its spectral and spatial descriptions of a scene. In most cases, hyperspectral images are used to obtain information from the captured space that cannot be easily acquired from standard color (RGB) images.

Hyperspectral and multispectral imaging have become valuable sources of information for several applications and topics, such as plastic recycling [11–13], food quality testing [14–16], agriculture [17,18], astronomy [19,20] and remote sensing [21–23].

The applications of hyperspectral technology to laboratory and field geological analysis have existed for no more than fifteen years. Before that, it was used exclusively for remote sensing applications, due mainly to equipment costs and size. Classical NIR-based spectrography methods [1], such as characterization of spectral absorption features and unmixing techniques [9], were the first approaches to be used, and a number of standardized methodologies were developed for mineral classification over geological samples.

Two milestones are well worth being mentioned in the mining industry context. The first one was the emergence of laboratory and field hyperspectral acquisition equipment, such as LabSpec [5–7], TerraSpec [24,25] and Hylogger [26] that are single pixel devices, and CoreScan [4] which is the most well-known multi-pixel laboratory platform. Additionally, there are diverse models of hyperspectral cameras [26]—most of them push broom style scanners—that are multi-pixel devices, for which the companies SPECIM [27–31] and Headwall [32,33] have established the industry standard. The second milestone was the development of software packages for hyperspectral data analysis—such as ENVI [34], and the geological analysis-oriented application The Spectral Geologist (TSG) [35,36] developed by CSIRO—that use classical spectrography approaches and statistical proprietary unmixing technique

developments (such as the iterated constrained endmembers method [8]) to perform semi-quantitative mineralogy on a single spectrum basis.

Several mining companies—BHP corporation at its operations in Chile (Minera Escondida and Minera Spence), for example—have taken hundreds of thousands of single pixel spectral samples from drill core composites using ASD tools [6,7] or similar, and analogous numbers of CoreScan [3] and HyLogger images from drill core trays have been taken. These images and spectra have been used to deliver mineralogical and geological data—mainly for variables such as alteration, clay content, lithology and mineral content—to build better geological and geometallurgical long and medium term models.

By the classical NIR spectrography approach, we mean methods that use each individual spectrum as a source of mineralogical or geochemical information. These methods are based on the identification and analysis of specific absorption features over the spectral range under study. As spectra are captured from mixed minerals, unmixing techniques [9] are used to decompose them nonlinearly into a set of real reference spectra, taken from pure mineral samples.

One of the major challenges of this approach is the need to consider representative spectra for pure mineral samples as a comparison base. The main issue with this method is that it does not consider the fact that the acquisition of hyperspectral images is a stochastic process. The phenomenon under study is the reflection of electromagnetic energy that (a) is emitted by sources that can vary and (b) is reflected by imperfect and highly complex surfaces. Then, engineers have to struggle to control acquisition conditions to generate the best references and images, trying to avoid absorption feature variations, which is eventually impossible.

Machine learning techniques have been widely implemented for the characterization of geological samples and free walls in mining deposits [37]. One example from remote sensing is the approach in which a classifier is independently applied to the spectrum obtained for each pixel and results are used to label the whole pixel set. Going beyond individual pixel information by understanding the spatial structure of a hyperspectral image is important to improve analysis results. High-dimensionality data usually decrease the performance of normal pixel-wise classifiers, and using segmentation maps to train support vector machine classifiers has proven to be beneficial [38].

In the analysis of the hyperspectral information of a source with spatial correlation, using information coming from its neighbors is very useful in classification, detection and other applications.

2.2. Geometallurgical Hyperspectral Analysis

For the large mining industry, being able to generate predictive control strategies for metallurgical processes is a structural improvement of great economic interest. Until recently, the main approach has been to build long and medium-term geometallurgical models [39] for ore deposits and then to track ore batches to the plant through production programs. The problem is that operational processes for mineral extraction and handling are very complex and influenced by multiple events and decisions, which makes it difficult for short term production planning to take advantage of geometallurgical models.

Once the mineral is stockpiled, and then crushed and heading towards the milling process lines, it is feasible to predict the route and time it will take at the milling or flotation processes. This can be done by using simulations and residence times calculations. To have a real time estimator of geometallurgical variables as input for predictive control strategies at this point would be a key element, but it does not exist today [40].

In this context, hyperspectral images taken from milling feed minerals or from composite minerals coming from drill core segments can be very useful for correlations with geometallurgical variables, to predict them in a real time (milling feed) or high resolution (ore deposit 3D model) basis.

2.3. Spectral Variability

Classic NIR or VNIR-SWIR spectrography recognizes minerals or chemical elements presence in spectral absorption features. Mixed minerals produce coexistence or superposition of these features along the spectra, and the more complex the mineral sample under study is, the more complex and denser is the feature set to be identified.

Underlying this classical feature characterization approach to mineral spectral analysis, there is the idea of having standardized spectra available. Therefore, considerable efforts are being made to establish appropriate laboratory conditions. It is expected then to be able to have a laboratory-level characterization of the black and white references, and thus to have perfect environmental conditions that allow transforming the spectra obtained from samples of interest into spectra that are free from the influence of any external phenomenon.

Each spectrum instance, taken from the same mineral sample, results in slightly different spectral data. Variability is unavoidably present in all aspects of sensing procedures. The question is which set of spectra better represents the phenomenon under study. One possible answer is that an ideal acquisition environment and fully controlled sample conditions will deliver the best representation. This conception leads us to the problem of having to reproduce that environment and those conditions in order to have a general standard.

Thus, even under highly ideal conditions, variability always arises. However, this variability within a set of spectra (captured from the same mineral sample) can yield to a probability distribution that can be thought of as the best sample spectral representation—although many individual spectra present different absorption features. Moreover, all spectra coming from a specific sample distribution will be correlated with each other. Attempting to work without this in-sample-self-correlation will likely produce statistical models that will fail to satisfy the sample element independence hypothesis. Then, when this spectra-statistical-characterization hypothesis is assumed, the first thing to do would be to test the variability of spectra under very controlled conditions.

Relevant efforts to overcome issues discussed above have done by generating robust methods to normalize spectra, using white and electronic noise (dark) references and then using spectra to apply feature extraction methods and machine learning techniques. We have used normalization approaches within one image with reasonable results. However, other issues arise when they are applied to a complete spectra data base, in which different groups of spectra have been gotten under different ambient conditions or with different pieces of equipment. Then, variability sources are diverse and difficult to avoid, or normalize, not only for images but also for white references.

An industry mineral sample standard acquisition setting for VNIR-SWIR data considers using white references that provide information about the variability due to ambient and illumination conditions. Certain variability ranges within white reference spectra are due in part to illumination changes produced by lamp temperature variations. We provide two examples to illustrate these issues in Section 6.1.

2.4. Spatial Context Capture

So far, we have described spectral variability as a constitutive element of hyperspectral images, and it has been suggested that this variability could be a very informative element as a proxy for geometallurgical characterization. In addition, spatial variability is another inherent element that is present in a HSI for a geological sample that delivers valuable contextual information [41–43]. The spatial distribution of the HSI pixels, as in any other kind of image, can be called texture [26,44,45].

Likewise, texture of a rock includes its size and the form and disposition of minerals that constitute it [41,42]. Texture indicates the building mode of the rock and describes relationships among its components which helps to identify the forms of minerals and their geometrical relationships. At mineral deposits, texture analysis delivers relevant information about structural and content properties, such as zoning, inclusions, twins, intergrowths among several minerals and the forms and size of grains [46].

When rocks are crushed as part of metallurgical processes, samples obtained are fine, which makes it more difficult to identify their textural characteristics. This entails greater difficulties to conclude on rock information by visual observation [47]. Even using microscopy, as mineral becomes finer it becomes more complex both to identify orientations or interactions between minerals and to describe the existence of an overlapping process [48]. Some minerals can be identified and some contact relationships between them can be determined, but not conclusively [47,49].

In hyperspectral imaging, texture can be seen from the pixel size point of view. If the sample is fine (powder) then we can have a number of grains in one pixel only. As a reference, typical experimental settings have pixels of around 250,000 square micrometers. In this case the combination of elements that produce one pixel spectrum is a different spectral mixture than the one produced by the rock surface, when the sample is coarser, and it is covered by groups of pixels [49,50]. Therefore, textural analysis in hyperspectral imaging makes sense when a set of pixels can cover a meaningful rock surface portion and loses importance when there is too much grain in the mixture [41,48].

However, textural analysis still has great potential to be used in the characterization of geological samples using hyperspectral images from a certain point of view. A very interesting perspective is related to grouping or clustering surface zones to be labeled by expert knowledge and then applying regression or classification algorithms that help geological mapping at different scales (face mapping, drill core mapping). In this context there are some filters for image processing that are capable of smoothing content while preserving transitions [51]. For example, bilateral filtering is a simple, non-iterative scheme for edge-preserving smoothing. By applying filters of this type on the spectral and spatial dimensions of a HSI it is possible to find relevant bands while allowing a strong dimensionality reduction. A possible approach is to first apply bilateral filtering, allowing HSI data to be converted into a smoothed version where transitions are preserved. The edge-preserving smoothing version of the HSI can be used to create a weighting of original bands. Then it is possible to define a spectral domain reduction constrained by the expected number of final bands. This allows one to transform the HSI data into a multispectral image (MSI). The bands in the resulting MSI can be estimated by the weighted linear combination of the HSI, conditioned by its bilateral filtered version [52]. This approach produces a MSI or even a RGB image that focuses on the texture and spatial content while strongly condensing the spectral information.

Image texture analysis has been well studied in recent decades in the context of computer vision [45,53]. We have performed some proof of concept in problems of texture classification and segmentation using image texture methods, including locally binary patterns [54], wavelet transform [55] and Gabor filters [56], among others.

We have studied different types of images that presented particular geological patterns in drill core samples, including aphanitic, phaneritic and porphyritic geological textures, and also some geological structural patterns such as stockwork and veinlets. Based on this problem, we created a rock image database and characterized it according to geological expert criteria [57,58] (see Figure 1).

A first method to perform texture classification using coefficient histograms of the wavelet transform sub-bands [59,60] was developed. Each of these sub-bands presents Gaussian behavior on the coefficient value distribution and can be modeled as a generalized Gaussian distribution (GGD) [61]. Finally, distribution parameters can be compared using statistical metrics [60] in each sub-band and then results can be fused in a metric to quantify similarity between both rock textures analyzed.

Another development in this line of research is the incorporation of cartoon-texture separation of image contents based on an optimization strategy that uses non-linear total variation methods [62,63] and features obtained with a multiscale shearlet transform for edge detection and texture characterization [64,65]. Then we proposed a chain sequential classification approach specifically designed for natural rock texture discrimination [57], which eases designing ad hoc features and provides a method to address the limited image data issue.

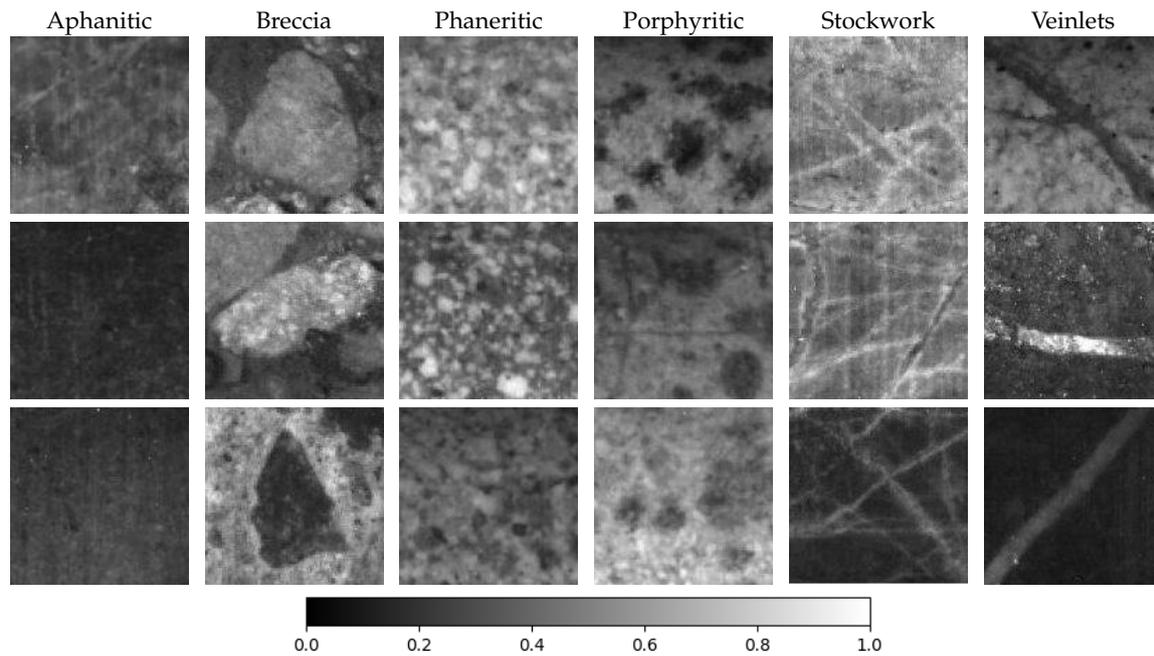


Figure 1. Examples from image rock texture database. Each class (presented in a separate column) was selected according to expert criteria, including geological structures and textures. Colormap: Images are presented in grayscale denoting for each class its specific visual characteristics (e.g., directionality, shape and sizes) which are described by its own spatial texture.

Finally, we explored using variogram-based descriptors for image textures analysis (see Figure 2). Variogram is frequently used in areas such as geostatistics and remote sensing for the study of phenomena expressed in a region that presents characteristic spatial continuity [66,67]. To describe this spatial continuity at different scales and in different directions, we proposed three variographic descriptors by using different patterns that characterize textural properties of images under study [58].

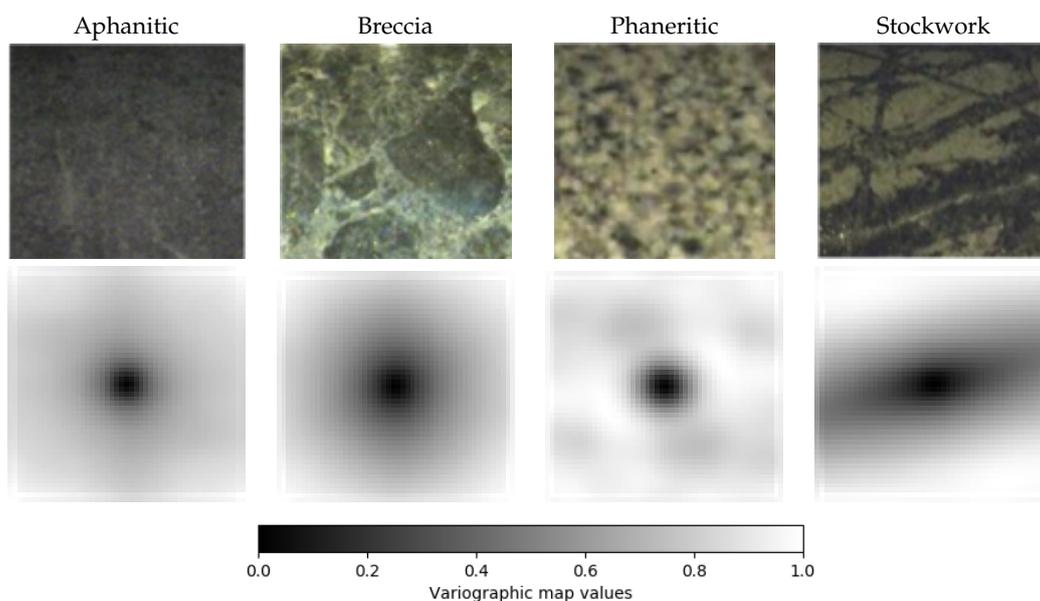


Figure 2. Original images in RGB color space from different classes of rock texture (**top**) and their corresponding variographic maps with colormap (**bottom**). Each variographic map (zoomed 2×) denotes the spatial continuity information present in the texture using normalized grayscale—from dark tones (low variogram values) to light tones (high variogram values).

2.5. Spectral Context Capture

Variability of a mineral sample surface, observed through spatial texture patterns, delivers information about geological or metallurgical sample features. Nevertheless, when considering an image in the visible range the simplest spectral information is color. In this case, the challenge is to isolate the feature that is exclusively associated with color, such as hue or chromaticity, and try to correlate it with a specific geological variable as mineralogy.

Previous work considers this approach and has been the theoretical foundation for our characterization approach described in Section 4. We developed a methodology based on the assumption that a single color, as a perceptual chromaticity phenomenon, cannot be considered as single point in a color space, but rather a whole statistical color distribution which summaries a complex combination of hue, light and shadow [68]. The color distribution is obtained using the technique of histogram partitioning to identify meaningful modes in the hue histogram [69,70] (Figure 3a), producing image color segmentation (Figure 3b)—assuming the image has been previously projected onto a hue-saturation-lightness (HSL) color space. Then every meaningful mode produces a unimodal histogram, which is considered a single color distribution that can be associated with a specific mineralogy. Subsequently, different mineralogy types are compared using both a statistical test and a statistical distance, like the Kullback–Leibler divergence.

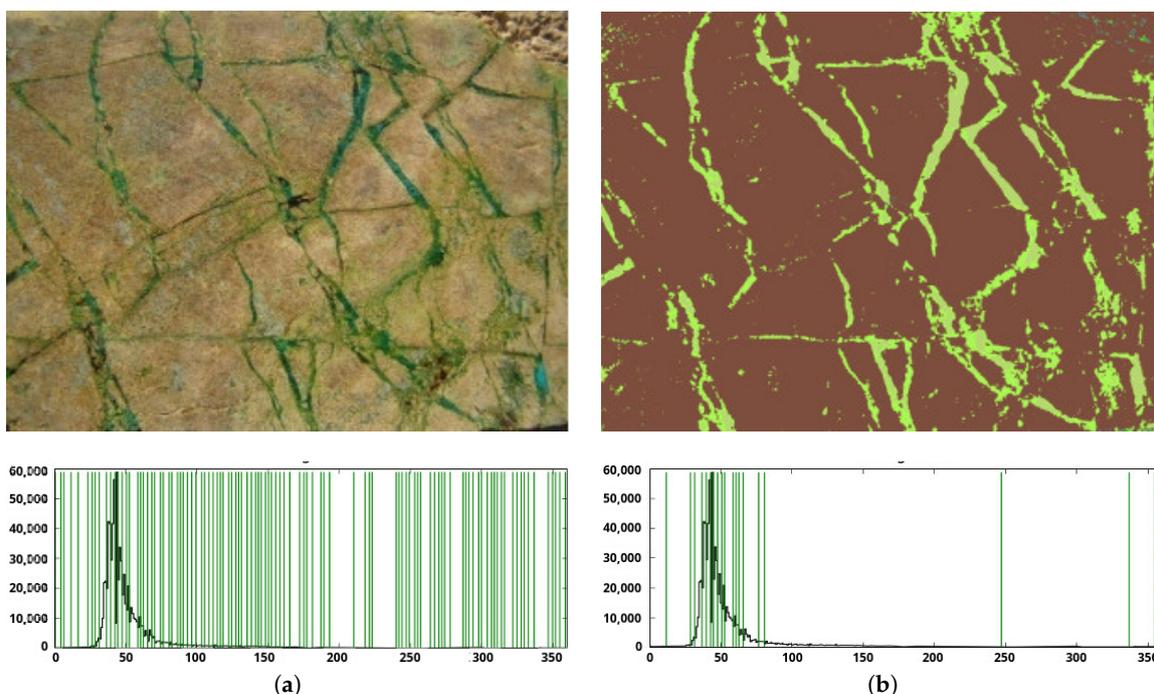


Figure 3. Segmented images in RGB space (**top**) and their corresponding histograms of hue values (**bottom**) with partitions. Column (**b**) show the results of color segmentation by searching unimodal distributions in the color histogram of the hue channel (column (**a**)).

2.6. Dimensionality Reduction

There are a number of feature extraction tools that have proven to be robust in many scientific and industrial areas. For example, some approaches seek to find bands or groups of bands that contain most of the energy of the original spectra [71], or they seek to estimate a regression that minimizes an error measure between the original spectra and the reduced version, such as partial least squares regression (PLS) [10,72,73]. It is also possible to take advantage of machine learning and deep learning techniques, such as convolutional autoencoders, which make it possible to find low-dimensional representations that preserve considerable amounts of information [74–76].

A versatile alternative for robust dimensionality reduction is using wavelet transform and filter bank pruning, which is based on this transform. There are several families and variants of wavelet transforms (such as Haar, Daubechies, Shearlets and others) that allow multiscale representations of spectra [59,77,78]. In the case of wavelet filter banks, it is possible to find components that are the most robust ones to characterize the type of data available. Nevertheless, most of these approaches are limited in big data or highly noisy data applications because required computing time increases considerably under these scenarios [79–81].

Various properties are sought in a robust, low-dimensional representation of data, among which the localization of scale, decorrelating power, completeness of representation, behavior in singularities or discontinuities and efficiency of the implementation stand out. Several works have proposed the use of wavelets as a method to reduce the dimensionality of hyperspectral image spectra [81,82], since it is possible to exploit the fact that their wavelet representation is sparse. Then, it is possible to combine the atomic decomposition and denoising tools to improve low-dimensional sparse representations [80].

2.7. Hyperspectral Image Clustering

Clustering methods attempt to find groups of similar objects, while dissimilar objects remain separated in different groups or isolated in a group for noisy data [83–85]. Objects can be spatial data points, feature vectors or patterns. Typical clustering techniques include centroid-based clustering [83,86], spectral clustering [87,88], density-based clustering [83] and others [89]. These techniques usually perform well for reduced noise and low dimensional data. In the case of complex shaped groups, which reside in high-dimensional spaces, classical techniques have insufficient performances because they are either deficient or lacking in noise and variability considerations, or limited by their assumptions on topological and spatial characterization in the feature space.

Over-segmentation can be used to create segmentation maps and perform spectral–spatial classification [90]. It is a good method for understanding spatial structure of an image in an unsupervised fashion, as resulting super-pixels can be later joined in order to search for the final targeted segmentation result of the HSI. Here, super-pixels are image regions, built from pixels that share common information and that are relatively close in the spatial dimension. An over-segmentation algorithm produces an overestimated number of super-pixels that might be joined together, because of their closeness, producing real segments [91,92].

Several methods can be found in literature for extracting super-pixels from an image [43,44,93–95]. In the specific case of HSI, techniques such as HSeg [96], partitional clustering [89,97] and Watershed [98], among others, could provide valid options for super-pixel clustering [44]. In the case of HSeg [96], the approach consists of an iterative form of region growing where each step reduces, by at least in one, the number of segments found. In each step the dissimilarity between segments is computed; the pair of segments that present the minimum value (T) are merged; and segments not contiguous with each other are merged if they are under a defined threshold.

Watershed approach has been applied for HSI segmentation [98]. For it, a gradient of every point in the image is computed. This new gradient space is then thresholded by an user proposed value, and the lower values in a neighborhood are clustered together. Segments are split, thanks to the higher gradients, which imply a frontier transition [38]. Furthermore, Watershed in HSI corresponds to a good segmentation method when there are relatively big structures which can be clustered inside the image [38]. Its downside is that gradients are not always easy to compute, and good measures in higher-dimensional spaces tend to be more resource demanding, because they require gradient choices, involving morphological transforms, such as Beucher gradient [38].

In addition, partition clustering methods have been successfully used to find segments inside HSI pixels. In this case, clustering is achieved by single partitions of the data into disjoint classes [99]. K-means and its modifications are the most common approaches to perform partition clustering for segmentation [85,86,100].

Aforementioned methods and other classical strategies, such as k-means, have some drawbacks [87,100]. Firstly, these approaches tend to be very inefficient when performed over a large set of data points, because in each iteration, distance of each pixel from all cluster centroids has to be computed and stored. Secondly, these methods are usually very sensitive to their parameters, which also depend greatly on the nature of data. Finally, k-means is designed to minimize a certain measure of distance between each cluster member, which does not take into account spatial information. Some works explored using spatial information in a modified version of k-means, which uses an alternative objective function to minimize a linear combination of spectral and spatial information [99].

Several works that provide insights about classical methods do not provide any comprehensive strategy of clustering that considers both noise or variability of data [101] and spatial structural aspects [99]. This consideration is demanded when working with high-dimensional phenomena that are typical for HSI data.

3. Related Work in Classical Strategies for Geological Spectral Characterization

Hyperspectral analysis is not a new topic in mining. A comprehensive review of close range, ground-based hyperspectral imaging has been recently presented for mining applications at several scales in [102]. Here, several cases of single spectrum mineral identification are described. This approach differs from our method because it does not consider spatial context or spectral variability.

Most common classical approaches for HSI work address inherent variability of spectra by using continuum removal correction [103–106] to ease comparisons among several spectra [107]. Central idea behind continuum removal is to clean or normalize data set to eliminate disturbances—as data are sensitive to diverse factors, including non-uniform thermal properties, mineral background reflection, acquisition angle and several other parameters. While this approach facilitates continuous removal of components in the spectra and highlights present absorption bands, it cannot not be applied in the proposed pipeline, as the latter uses complete spectral information for the inference process. In this context, a continuous spectrum provides a useful source of information for characterizing the spectral variability and allows one to detect spectral clusters. If there is spurious information (or noise), it is absorbed by the dimensionality reduction and the autocorrelation characterization of hierarchical model. Moreover, proposed method is not intended to extract any deterministic features from spectra, as spectral information bands are always regarded as predictive variables to be associated with any geometallurgical interest variable.

An interesting comparison of clustering and classification tools for automatic mineral identification has been presented in the literature using long wave infrared HSI data [107]. Here, a low-rank sparse principal component analysis (Sparse-PCA) approach was compared to classical spectral angle mapper (SAM), spectral information divergence (SID) and normalized cross correlation (NCC). The latter were used to extract features from the spectrum in the form of a false color composite. Low-rank sparse-PCA allowed them to extract spectral references which showed high similarity to the ASTER (JPL/NASA) spectral library. These tools compare individual spectra against a standard library, and performance improvements among them are not conclusive from the experiments provided.

A commonly visited approach for spectral analysis attempts to find pure spectra (called endmembers) which form a base to properly describe all available spectra. The main idea here is to infer contributions from all pure components present in a complex spectrum to build a matrix of individual pure spectra and corresponding concentration profiles which best fit experimental data. An interesting example of this technique is shown in Pisapia et al. [108], where tracking of organic carbon hidden in rocks has been explored by integrating chemometrics and a specific type of spectral data (synchrotron radiation-based Fourier-transform infrared microspectroscopy (S-FTIR)). The analysis has been interestingly performed using a combination of principal component analyses (PCA) and the mixed approach called multivariate curve resolution—alternating least squares (MCR-ALS).

Another very interesting work that combines dimensionality reduction and statistical analysis for mineral classification problem was presented by Ekanayake et al. [109]. This work attempts to

obtain ilmenite content mapping for a real case by using principal component analysis (PCA) as a dimensionality reduction strategy. Then, a Euclidean distance-based method is used to extract pixels containing soil, and classical Fisher discriminant analysis (FDA) is used to split the data set into two distinct classes—ilmenite and non-ilmenite. A clever additional stage is proposed to correct the clustering result by using a probability density function—designed upon the mineral spatial distribution—which is very interesting because it allows one to incorporate expert criteria, in the same way as a kernel-trick distance does in the proposed clustering method, as described in 5.2. However, PCA reduction here produces high-dimensionality data variability expressiveness lost. The proposed method uses dimensionality reduction just to reduce computational cost—where energy preservation is intended to keep data variability and hence information.

Finally, an evaluation method for mineral identification algorithms has been presented by Yousefi et al. in [110]. This work analyzes the reliability of longwave infrared (LWIR) hyperspectral imagery mineral identification systems by considering the ASTER spectral library of JPL/NASA ground truth data, spectral angle mapper (SAM) and normalized cross-correlation (NCC) to create false color maps. K-means clustering is used to group mineral regions in different categories by working on transformed data using hue-saturation-value (HSV) principal component analysis (PCA) producing false-color composites. Here, there is again the assumption that there exists an ideal real spectrum set to be compared with by using the SAM metric, which is intended to focus on spectral features. Conversely, the proposed hypothesis relies on the idea that isolating ideal spectral components in field conditions is practically impossible.

4. Proposed Stochastic Characterization of Metallurgical Samples

As mentioned in Section 2, diverse aspects have to be considered when developing a robust method to estimate geometallurgical variables from metallurgical sample hyperspectral images. One of them is related to the incident light spectrum and how it affects HSI data reflectance characterization. In addition, there are diverse environmental aspects that affect the acquisition process and that are unavoidable sources of noise and variability.

We propose a stochastic characterization method which consists of process pipelines to train and predict/estimate interest variables from metallurgical sample hyperspectral images. Based on concepts of spectral variability, spatial and spectral context capture, hyperspectral data dimensionality reduction and clustering and hierarchical stochastic regression modeling, the proposed pipelines represent a new strategy for rapid geometallurgical estimation.

4.1. Working Hypothesis

Under homogenization conditions in the metallurgical sample, spatial context from original rocks is lost, and as shown in Section 2.5, color information from a RGB image can be used as a distinguishing variable to be correlated to mineralogical characterization. The extrapolation of this principle to HSI data is the core component of the proposed methodology. Single pixel (spectrum) characterization and spatial-spectral distribution along the sample surface are related to the variability of geochemical and mineralogical properties of the metallurgical sample.

Our hypothesis is that any relevant geometallurgical feature that needs to be estimated from a hyperspectral image taken from a metallurgical sample, resides on the spectral and spatial variability of the surface under study.

In this regard, we consider spectral data as samples drawn from statistical distributions—in the same way as in Section 2.5 we considered color as a statistical distribution in a color space. Thus, spectra are treated as groups of signals with noise, instead of considering them as single objects that need to be *rectified* in order to be compared to a reference spectrum coming from a standard database. Then, characterizing variability and dynamics of several spectra from pixels within a metallurgical sample allows direct training and correlations between this information and variables measured from the sample with standardized laboratory analysis or sensing methods.

4.2. Proposed Methodology

To test our hypothesis we have designed a new schema to estimate geometallurgical variables using hyperspectral images that is based on principles described in Section 2. Processing pipelines include three main tools. Namely, an energy-preserving dimensionality reduction method based on wavelets and the kernel trick, a hybrid spatial spectral iterative clustering method and a stochastic hierarchical regression model. In the following sections, processing pipelines are described in general terms, and then each tool mentioned is described in detail.

4.2.1. Training Pipeline

For this pipeline we will define each model observation as a spectral patch which is composed of a set of spectra, with or without spatial correlation, that are taken from the same metallurgical sample's hyperspectral image. These patches represent the spectral variability for different portions of the image/sample, as described in Section 6.

The first step in this pipeline is to characterize a number of metallurgical samples (M) using any standardized laboratory analysis for geochemical, mineralogical or geometallurgical variable vector determination.

Then, the proposed model training requires one to:

1. Prepare the M metallurgical samples considering to expose their surface as much as possible.
2. Acquire hyperspectral images from each metallurgical sample and to normalize them with respect to white and dark references to obtain reflectance images.
3. Reduce pixel spectra dimensionality to 10 percent of their original bands number.
4. Obtain an arbitrary amount of pixel patches from each image—25, 50 or 100 pixels for each patch would be a reasonable number, depending on the image and pixel size, and mineral sample grain profile.
5. Apply a clustering process to all patches from every image belonging to the training set. A patch is a metaspectrum (concatenation) which codifies their energy. The clustering is applied to an amount of clusters that has been determined by a sensibility analysis based on the specific data set under study.
6. Train a multilevel regression method for each characterized variable. Here, the base level is the actual patches set and the next one is composed of the patch clusters. Then, each patch has a label with the cluster it belongs to.

Thus, we obtain a hierarchical or multilevel regression model for each variable that contains the characterization vector from laboratory analysis. Figure 4 shows the entire process flow.

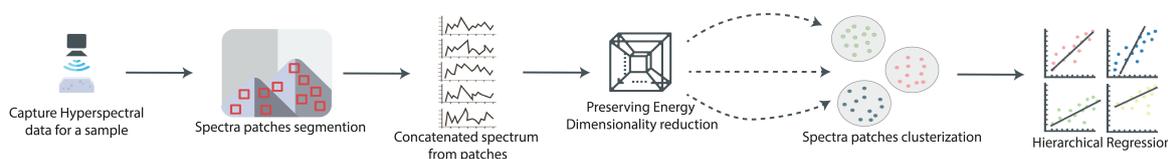


Figure 4. Hyperspectral image analysis by coherent clustering and hierarchical regression.

4.2.2. Prediction Pipeline

To estimate interest variables for a new metallurgical sample (not included in the training set) it is necessary to:

1. Prepare the new metallurgical sample for the acquisition by exposing its surface as much as possible.
2. (a) Acquire a hyperspectral image from the new sample, (b) obtain the reflectance image, (c) reduce dimensionality and (d) divide it into obtain a set of pixel patches, as many as possible, because the aim is to obtain an experimental metaspectral distribution at this stage.

3. Obtain and assign the closest trained cluster for each patch using the Hausdorff distance. If some patch is close enough to more than one cluster (according to some predefined threshold), then it must be duplicated and assigned to every close cluster to augment the number of elements for the *a posteriori* estimation.
4. Apply the regression functions (associated with the assigned closest cluster) to each patch.
5. Use regression function outputs to obtain an experimental distribution for each variable of interest.
6. Estimate each variable of interest using its experimental distribution. Expected value can be a reasonably good estimator for most cases, but if the experimental distribution is too complex (multi modal, for example) it could be a kind of naive estimator.

Figure 5 shows the entire estimation pipeline.

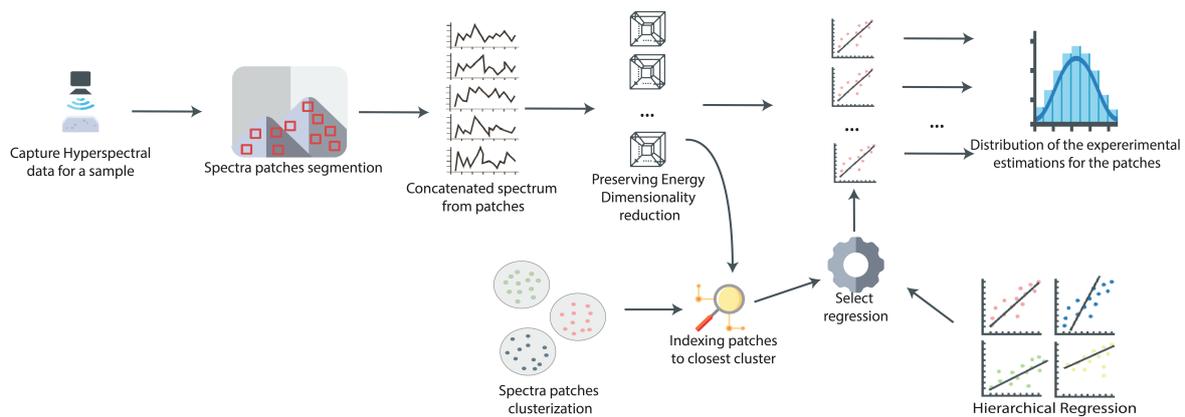


Figure 5. Prediction pipeline for a new sample.

5. Proposed Core Pipeline Tools

The following sections present the strategies studied and implemented for the main components of the proposed pipelines in Section 4.2.

5.1. Tool I: Preserved Energy Dimensionality Reduction

A common problem when dealing with hyperspectral images is data high dimensionality [111] (details in Section 5.2). There are various strategies that have been applied at diverse areas to manage this situation [112]. For example, it is possible to find methods for reducing data dimensionality by estimating features that can appropriately describe their contents. It is relevant to note that in the case of HSI data, near bands generally have a high level of correlation and acquisition process includes phenomena of integration in both spatial and spectral domains [113,114].

For HSI data, a robust approach must consider identification of features that have less dimensionality than input data, while preserving spectral and spatial content structure [115]. Furthermore, in our context, it is necessary to find a strategy that can make the reduced data suitable for Algorithm 1; i.e., reduced data must also be coding the original spectrum energy. Furthermore, method needs to be fast enough to be incorporated into processing pipeline (Figure 4).

Wavelet Approach

In the context presented in Section 2.6, Daubechies wavelets could be an useful alternative, as they have the shortest possible support for the given number of vanishing moments and an extreme phase [77,78]. Considering the compactly supported Daubechies wavelets [101] and a threshold for the scaling level j_0 (the scaling is indexed by j in this notation), a spectrum $\mathbf{x} \in \mathbb{R}^n$ can be transformed into a new feature space describing wavelet and scaling components of the signal. Using notation proposed

by Chen [79], original data can be represented by wavelet coefficients $c_{j_0,k}$ and scaling coefficients $d_{j,k}$ in the feature space, determined by orthonormal basis $\varphi_{j_0,k}$ and $\psi_{j,k}$ respectively.

$$\mathbf{x} = \sum_k c_{j_0,k} \varphi_{j_0,k} + \sum_{j>j_0} \sum_k d_{j,k} \psi_{j,k} \quad (1)$$

It is worth mentioning that scale coefficients represent spectrum (signal), after low-pass filtering, that preserves their smooth shape, while wavelet coefficients represent signal after high-pass filtering.

Once the transformation given by expression (1) is obtained, it is possible to preserve a high value scaling coefficient from wavelet representation only, considerably reducing spectrum dimensionality. Preserved coefficients can be used to reconstruct approximated spectrum $\hat{\mathbf{x}}$ and correspond to a sparse representation in wavelet domain, usually called *B-term representation* [116]—this approach has been used in others works to remove noise for highly noisy data [79].

Transformation for *B-term representation* can be considered as limit-sparseness processing of signal. The objective function of this sparse representation is:

$$\min \|\mathbf{x} - \sum_{j>j_0} \sum_k d_{j,k} \psi_{j,k}\|_p \quad \text{subject to} \quad \sum_{j>j_0} \sum_k \|d_{j,k}\|_0 = B \quad (2)$$

where $p \in [1, \infty]$.

Finally, for an integer B describing the number of preserved coefficients, it is possible to simplify notation of *B-term representation* by redefining its coefficients as $\{f_i\}_{i=1}^B$, $f_i \in \mathbb{R}$ and denoting $\bar{\psi}_i$ as the corresponding wavelet scaling basis for f_i coefficient. Thus, spectrum *B-term representation* is:

$$\hat{\mathbf{x}} = \sum_i^B f_i \bar{\psi}_i \quad (3)$$

5.2. Tool II: Hybrid Spatial Spectral Iterative Clustering

5.2.1. Clustering Method Description

Figure 4 shows that a clustering strategy is at the core of the proposed method. In the same way that meaningful mode extraction (Section 2.5) aimed to find the color distribution (i.e., all points in the color circle H when using a HSL color space) that represents a visually distinguishable color, the objective here is to group spectra so that they can be associated with a response or target variable value. Although any clustering algorithm could be a good starting point, we have to take into account that there are two dimensions that must be considered at the same time: the spatial domain and the spectral domain. In this regard, a modified version of the simple linear iterative clustering method (SLIC) [95] is presented.

SLIC method is an image segmentation strategy, based on a local spatial partition clustering approach, that has been widely adopted in applications involving common digital images, because it has shown: (a) a good performance, (b) to be fast at finding object borders and (c) a linear convergence time proportional to the number of target super-pixels and to the number of pixels contained in the image. The original SLIC strategy considers the image to be converted to the CIELAB color space [117] rather than using the traditional RGB space [95]. This is because it relies on color space perceptual uniformity, which in practical terms means that color representation needs to have some sort of information on signal energy—coded in CIELAB lightness channel—which makes it suitable to be modified for HSI, because spectral information can be regarded as HSI color representation directly.

Thus, proposed hybrid spatial spectral iterative clustering (HySSIC) method is straightforward and consists of a localized application of *k-means clustering*, using a distance measure that considers a combination of both spectral and spatial information. Algorithm 1 shows the overall method outline. An initial uniform grid of step s is created. Step s is calculated from the parameter $k \in [1, R]$ where R is

HSI spatial resolution—i.e., HSI total amount of pixels. After initialization, all cluster center positions are perturbed to guarantee convergence. New positions are such that they are the lowest gradient location in a $n \times n$ neighborhood computed as follows:

$$G_{i,j} = \|\mathbf{x}_{i+1,j} - \mathbf{x}_{i,j}\|_2^2 + \|\mathbf{x}_{i,j+1} - \mathbf{x}_{i,j}\|_2^2 \tag{4}$$

where $\mathbf{x}_{i,j}$ is the spectrum at location (i, j) , as described in Figure 6. Once all cluster centers are obtained, distance from every pixel in the HSI to each cluster center pixel is calculated using an ad hoc distance function:

$$d(\mathbf{x}, \mathbf{y}) = d_S(\mathbf{x}, \mathbf{y}) + \frac{d_R(\mathbf{x}, \mathbf{y})m^2}{s^2} \tag{5}$$

Expression (5) represents the concept of combining a spectral distance $d_S(\mathbf{x}, \mathbf{y})$ with a spatial distance $d_R(\mathbf{x}, \mathbf{y})$. Note that when $m = 1$, this combination depends on the uniform grid step s only—that is to say, if k increases, the spectral distance component importance increases. In any case, parameter m can control relative contribution of the spectral and the spatial distances. After all these distances $d(\mathbf{x}, \mathbf{y})$ are obtained, every pixel is assigned to the nearest cluster. This means that cluster centers need to be recalculated considering the average location and spectrum of all pixels assigned to the cluster. Then error updating operation is done using spectral distance only. Finally, the algorithm stops when error is less than threshold defined by parameter τ .

Algorithm 1 The HySSIC method.

```

procedure HYSSIC(HSI,  $k, m, \tau, n$ )
     $s \leftarrow \lfloor \sqrt{\frac{R}{k}} \rfloor$  ▷ Where  $R$  is the HSI spatial resolution
    Initialize a uniform grid (with step  $s$ ) of cluster centers
    Perturb cluster centers to the lowest gradient locations ▷ In a  $n \times n$  neighborhood
     $E \leftarrow \infty$ 
    while  $E \geq \tau$  do
        for each cluster do
            for each pixel in the cluster do
                Calculate the distance to the cluster center ▷ Using expression (5)
            end for
        end for
        Assign all pixels to the nearest cluster
        Compute the new cluster centers
         $E \leftarrow \{\text{Spectral distance between new and previous cluster centers}\}$ 
    end while
end procedure
    
```

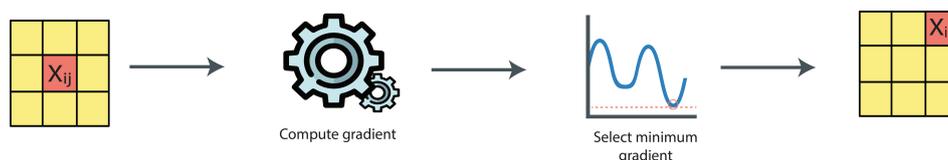


Figure 6. Procedure of the minimum gradient reallocation of a single centroid using 3×3 neighborhood hyper spectral data.

5.2.2. Spectral Distances

Spatial distance $d_R(\mathbf{x}, \mathbf{y})$ can be any metric defined in the pixel location coordinates context, such as the standard L_2 norm. However, it is in the spectral distance $d_S(\mathbf{x}, \mathbf{y})$ that we need to pay closer attention to, in order to take advantage of spectral information as much as possible.

Spectral Angle Mapper Distance

If the analysis has to be focused on spectral absorption features, the spectral angle mapper (SAM) [9] is one of most used distances. It has the nice property that it ranges from 0 to $\frac{\pi}{2}$ and geometrically can be seen as the angle between two vectors, computed using euclidean projection of one onto the other:

$$d_S(\mathbf{x}, \mathbf{y}) \equiv d_{\Theta}(\mathbf{x}, \mathbf{y}) = \arccos \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2} \right) \quad (6)$$

This measure is ideal to compare spectra where some spectral feature (like an absorption band) is expected to be present in the same band but whose magnitude is not necessarily the same on all spectra—which makes it similar to a cross correlation distance. For two spectra, SAM distance takes value 0 when a spectral feature is present in the same band with exactly the same magnitude in both cases, and increases as a spectral feature is present in the same band, but differ in their magnitude, reaching the value $\frac{\pi}{2}$ when a spectral feature is present in just one of the two spectra.

The main advantage of using SAM distance within expression (5) is that it provides the HySSIC method with some kind of interpretability in the classical sense. For instance, if $m = 0$, when Algorithm 1 stops, we obtain a group of clusters whose members share some specific spectral features.

Nevertheless, SAM distance is very sensitive when samples spectral resolution is high, which is the case for most HSI. This means that small differences between two spectra (i.e., when they are sharing a spectral feature to a higher extent) can still yield values close to $\frac{\pi}{2}$. This impacts directly one of key aspects of any cluster algorithm, which is sample separability. Thus, in higher spectral resolutions, numeric precision needed to compare SAM values is higher, leading the clustering method to fail due to erroneous metric information. Then, this spectral distance can be considered as an interesting starting point for $d_S(\mathbf{x}, \mathbf{y})$ in the HySSIC method, to gain some indications about samples, but it should be used with caution.

Kernel-Trick Distances

Regarding data (or spectra in our case) separability, one of the key to success of kernel methods is a mathematical tool called the kernel trick [118,119]. In short, the kernel trick consists of projecting spectra onto an extremely high-dimensional space, where both similarities and differences among them will be emphasized, ensuring their separability. If the target space is a reproducing kernel Hilbert space (RKHS), all its topological properties are captured by a unique function (associated with that specific RKHS) called *kernel*, which takes values on the original data space. This means that any distance construction can be made on the original data space, using the target space unique kernel, to capture the desired topology.

Thus, in our case, given two spectra (\mathbf{x}, \mathbf{y}) and a kernel function k , a spectral distance can be defined as a kernel-trick distance [100,120–122] as follows:

$$d_S(\mathbf{x}, \mathbf{y}) \equiv d_k(\mathbf{x}, \mathbf{y}) = \|\Phi(\mathbf{x}) - \Phi(\mathbf{y})\| = \sqrt{k(\mathbf{x}, \mathbf{x}) - 2k(\mathbf{x}, \mathbf{y}) + k(\mathbf{y}, \mathbf{y})} \quad (7)$$

The target space projection function Φ is shown just to make the implicit projection clear, as the HySSIC method considers the part that only involves kernel k . An appropriate choice of this kernel may result in a big enhancement of the clustering performance, as it may change the topological properties of target space where spectra are assumed to be. This impacts directly the way spectra are separated and how different spectral features are considered for it. Moreover, in the same way as covariance functions in variogram analysis, kernels can be operated among them to obtain more complex target spaces, providing an interesting field of spectral clustering analysis. Some of the basic kernels that can be used within the HySSIC method are:

- $k(\mathbf{x}, \mathbf{y}, \alpha, \beta) = |\alpha \mathbf{x} \cdot \mathbf{y} + \beta|$ (linear)
- $k(\mathbf{x}, \mathbf{y}, \alpha, \beta, d) = |\alpha \mathbf{x} \cdot \mathbf{y} + \beta|^d$ (polynomial)

- $k(\mathbf{x}, \mathbf{y}, \theta) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|}{\theta}\right)$ (exponential)
- $k(\mathbf{x}, \mathbf{y}) = \sum_i \min(x_i, y_i)$ (histogram intersection)
- $k(\mathbf{x}, \mathbf{y}, a) = \prod_i \cos\left(1.75\left(\frac{x_i-y_i}{a}\right)\right) \exp\left(-\left(\frac{x_i-y_i}{\sqrt{2}a}\right)^2\right)$ (wavelet)

Note how every kernel shows different properties for the spectral distance.

5.3. HySSIC Examples

As an illustration of impact on cluster distribution provided by the proposed HySSIC method using the SAM spectral distance, resulting segmented 2D images for two samples are described in Figure 7. Here a sample is presented that corresponds to radial cuts of drill core sample sections. As these samples preserved the spatial distribution of the original rock, the HySSIC methodology was applied with a high value for m , prioritizing formation of clusters that spatially segment available information. As mentioned above, HySSIC method allows adjusting the trade off between spatial cohesion and relevance of spectral content, by giving different values for parameter m , letting the size of spatially located clusters be adjusted on a large scale. This type of spatial grouping is useful, for example, for assisting labeling process in training sets for automated classification algorithms application.

In the case of manual labeling tasks performed by an expert, the preservation of spatial content and spectral characterization allows useful clusters to identify spectral signatures of minerals or materials of interest. HySSIC method first defines clusters like those presented in Figure 7 (left), but they can then be grouped in an agglomerated way (Figure 7, right), by decreasing the value for m , to be able to match segments that are spatially separated, but that share similar spectral contents.

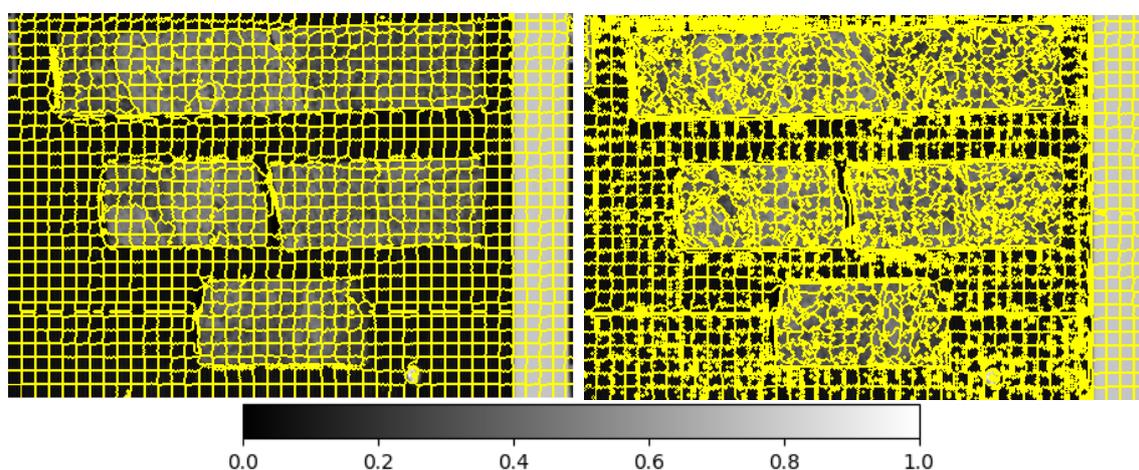


Figure 7. HySSIC applied to samples with spatial structures that must be preserved. **Left:** prioritizing spatial cohesion. **Right:** increasing relevance of spectral content. Colormap: the colors black and gray describe background and objects in the image respectively; yellow lines describe the boundaries of the clusters.

The proposed HySSIC method includes a posterior agglomerate clustering strategy, with lower values for m , that associates previously estimated clusters, in order to manage samples where spatial structure is not present, such as fine (powder) material. As an example, Figure 8 describes samples of soil saturated by different levels of oil.

For these samples, zones with similar spectral content may be spatially separated. In the initial application of HySSIC method, clustering with spatial cohesion is found, but in the final stage these are agglomerated in such a way that robust groups of spectra are characterized both in spatial and spectral domain. In this illustrative example, it is possible to identify areas that share similar doped oil content. By construction, from left to right, for each sand sample, the oil saturation level was gradually

increased. It is consistent with the final formation of agglomerated clusters proposed by the method. In the upper section of Figure 8, pixels (spectra) that have been assigned to the same agglomerated cluster are seen in the same color, where only the agglomerated clusters related with doped areas are presented. Lower section shows all the clusters proposed by the HySSIC method before agglomeration. Lower section shows all the clusters proposed by the HySSIC method before agglomeration.

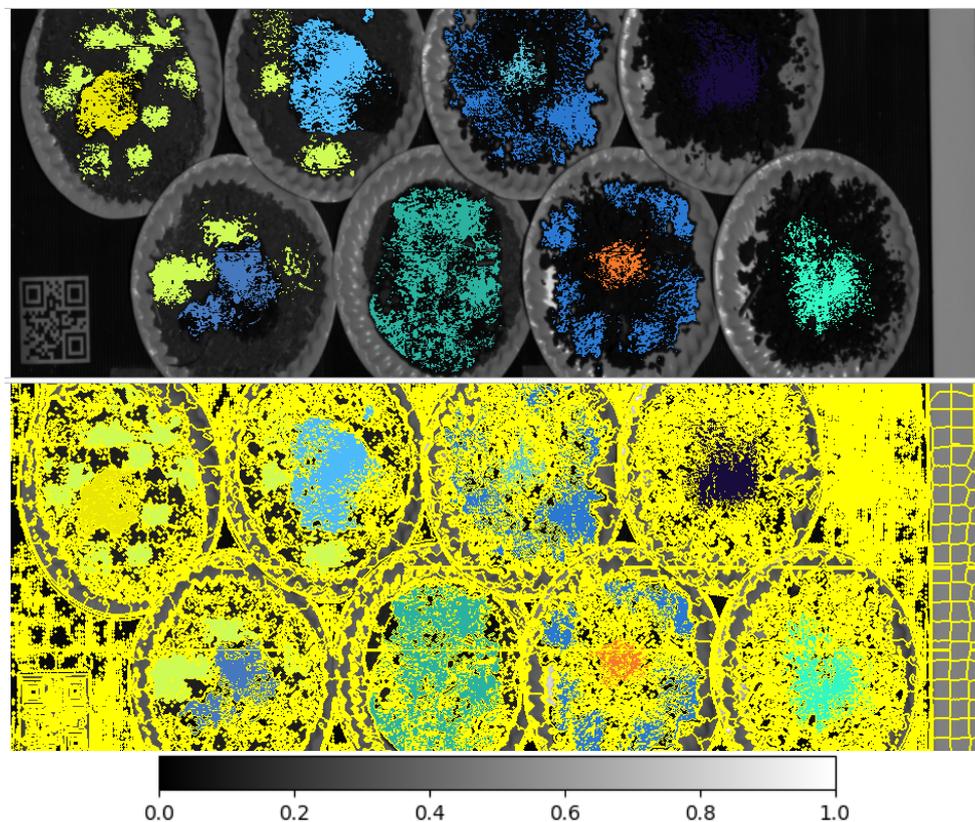


Figure 8. HySSIC applied to oil-doped soil samples. Top: clusters for doped areas. Down: full HySSIC clustering output. Colormap: black and gray colors describe background and objects in the image; yellow lines describe the boundaries of the clusters. The remaining colors define pixels sharing the same cluster.

5.4. Tool III: Stochastic Hierarchical Regression Model

Classical NIR analysis based on multi or hyperspectral data attempts to estimate quantitative or semi-quantitative measurements of mineral species of interest, to establish their correlations with metallurgical variables. Our proposed approach considers the spectral response variability of a set of pixels (spectra) from a metallurgical sample to be directly correlated with variables of geometallurgical interest by modeling how the latter is associated with different clusters of samples. Thus, in this section, we assume that a proper clustering scheme, such as the proposed HySSIC method, has been used previously on the data.

5.4.1. Mixed Effects Statistical Modeling

In the search for a framework that allows defining a hierarchical model that takes into account different geometallurgical samples variability, it is possible to use a linear mixed model (LMM) [123–125] that allows accounting for clustered data.

LMMs, also called *random effects* models, are described as statistical models in which some parameters, called *effects* in this type of formalization, exhibit some form of random variation. They are an extension of simple linear (fixed effects) models where variations in observed variables are described by systematic components (or fixed effects) and non-systematic components (or random effects). Its hierarchical nature arises when noticing that sample clusters represent a second data organization level starting

from the whole sample as the initial level. Thus, from a hierarchical point of view in LMM, fixed effects represent the behavior of the initial level (the whole sample), while the random effects represent variations in the second level, i.e., within or between clusters. Hence, LMMs are a general class of multilevel statistical models that can be used to model dependent data. It is also worth mentioning that in LMMs it is also possible to estimate the amount of variability of the random effects, which for example could describe relationships that contribute with valuable information on the processes under study.

More formally, if we denote the outcomes by \mathbf{y} , we can assume that its sampling distribution is given by the LMM (generative process):

$$\mathbf{y}|\mathbf{X}, \mathbf{U} = \mathbf{X}\beta + \mathbf{U}\gamma + \varepsilon \quad (8)$$

with

$$\begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \underbrace{\begin{pmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}}_{\Sigma} \right) \quad (9)$$

Here, \mathbf{X} and \mathbf{U} are the design matrices. Matrix \mathbf{X} contains the factors with fixed effects that we want to study and β is the vector that denotes these effects. Factors matrix \mathbf{U} , with random effects γ , contribute to the variability in $\mathbf{y}|\mathbf{U}$ and contains the cluster assignments for the observations (\mathbf{y}, \mathbf{X}) . Terms \mathbf{G} and \mathbf{R} are block diagonal covariance matrices whose simplest forms are multiples of identity. Both fixed and random effects β and γ respectively, which in principle are unknown, can be obtained by maximizing the log-likelihood function given by the joint density $f(\mathbf{y}, \gamma)$ —note that with this model, $\mathbf{y} \sim \mathcal{N}(\mathbf{X}\beta, \Sigma)$, $\gamma \sim \mathcal{N}(\mathbf{0}, \mathbf{G})$, $\varepsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ and $\text{Cov}(\gamma, \varepsilon) = 0$.

5.4.2. Correlation-in-Samples Modeling

Now that we have a working strategy, defining the model elements in the geometallurgical HSI analysis is straightforward. Geometallurgical variable outcomes will be coded in the vector \mathbf{y} . Spectral information will be coded in the fixed effect factors design matrix \mathbf{X} . As every individual spectrum may be contaminated or not fully representative of a sample under study, it is assumed that groups of spectra are correlated among them. It is then proposed to group spectra as a base unit that describes the variability of the sample content by clustering them with the proposed HySSIC method, using a kernel-trick distance to ensure spectra separability. These units will be coded as a random effect factors matrix \mathbf{U} in our model. Based on our ideas and assumptions, these units will have an internal variability and some dynamics that are appreciable in their spectral distribution, and that are expected to change for different samples and for changes in materials of interest contents or serve as a proxy for the geometallurgical variable outcomes.

6. Experiments

6.1. Spectral Variability in White References Examples

Here we present two experiments to illustrate issues introduced in Section 2.3. First one was done to evaluate illumination effects changes during single pixel hyperspectral acquisition process over samples under standard laboratory conditions. Spectral images of white references were taken every 5 min, during one hour. The idea was to understand whether illumination changes can be interpolated or approximated by improving white references usage according to the working protocols. In the second example, white references were recorded every 30 s during an hour of the acquisition system operation. A variability analysis is shown as result.

6.1.1. Variability of Ambient Conditions

We present an example of variability in the spectral domain for a standard white reference set. The spectra were obtained under laboratory conditions, and the process is intended for geological sample spectra acquisition in a mining operation. The 11 spectra used were a subset of 110 of them taken every 5 min during an acquisition session. The hyperspectral instrument LabSpec from ASD, used in this example, has wavelength bands from 350 nm to 2500 nm (with spectral resolution of one nm), where the information from 3 sensors is merged to generate the final spectrum. It is possible to appreciate that a stable response can be achieved in the target reference (see Figure 9). However, as shown in Figure 9, even under strictest laboratory conditions it is possible to see small disturbances. When the objective is precise reconstruction of spectral features, these alterations can lead to errors in the estimation or identification of interest species.

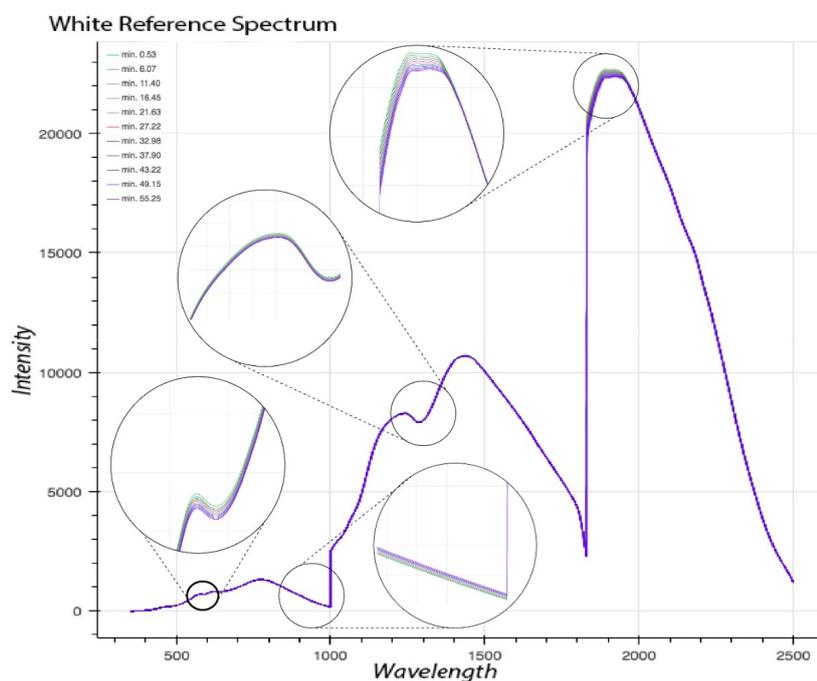


Figure 9. Global view of the 11 spectra of the available white references.

6.1.2. Inherent Variability

We have carried out a demonstration experiment with our hyperspectral cameras and a white reference material called *Optopolymer* which is characterized by a very stable response along the spectrum (See Figure 10). One camera comprising from 400 to 1000 nm of electromagnetic spectrum (SPECIM device; model Andor Zyla sCMOS with 488 spectral bands), and another covering from 1000 to 2500 nm (SPECIM device; model SWIR MCT with 288 spectral bands) were used. In this case white reference reflectance response characterization spectra are presented separately for each camera. In the following sections, examples images taken by each camera are merged in both the spatial and spectral domain, obtaining spectra in the range of 400 to 2500 nm. Tools based on *SimpleITK* library for python [126] have been considered for spatial fusion, while concatenation of disjoint spectra has been applied for the spectral fusion.

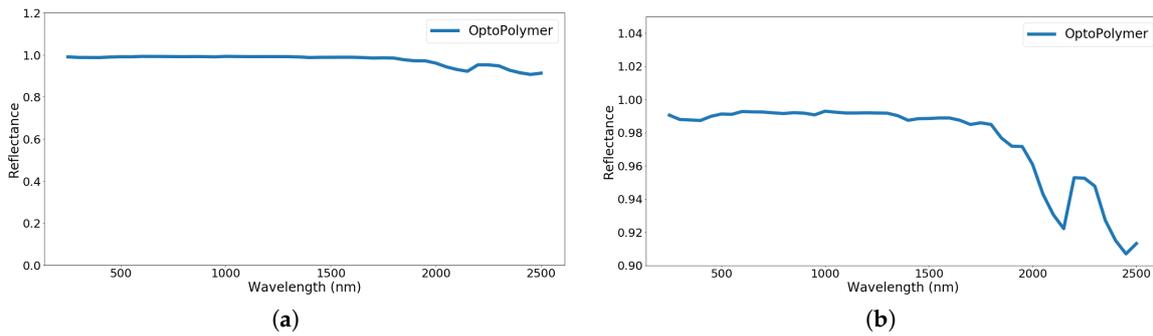


Figure 10. (a) Optopolymer reflectance response. (b) Optopolymer reflectance zoomed in.

Lighting conditions have been fixed and stabilized using halogen lamps operated with continuous switched power sources. In this scenario scan lines from the same section of the target reference have been acquired at a rate of 45 Hz for a period of 30 s. This process has been repeated every 10 min for an interval of 90 min.

From Figures 11 and 12, it is possible to appreciate that even when stabilizing the external conditions strongly the spectral response continues to be affected by small out-of-control disturbances and by the sensor dynamics (see Figures 11b and 12b for dark reference dynamics).

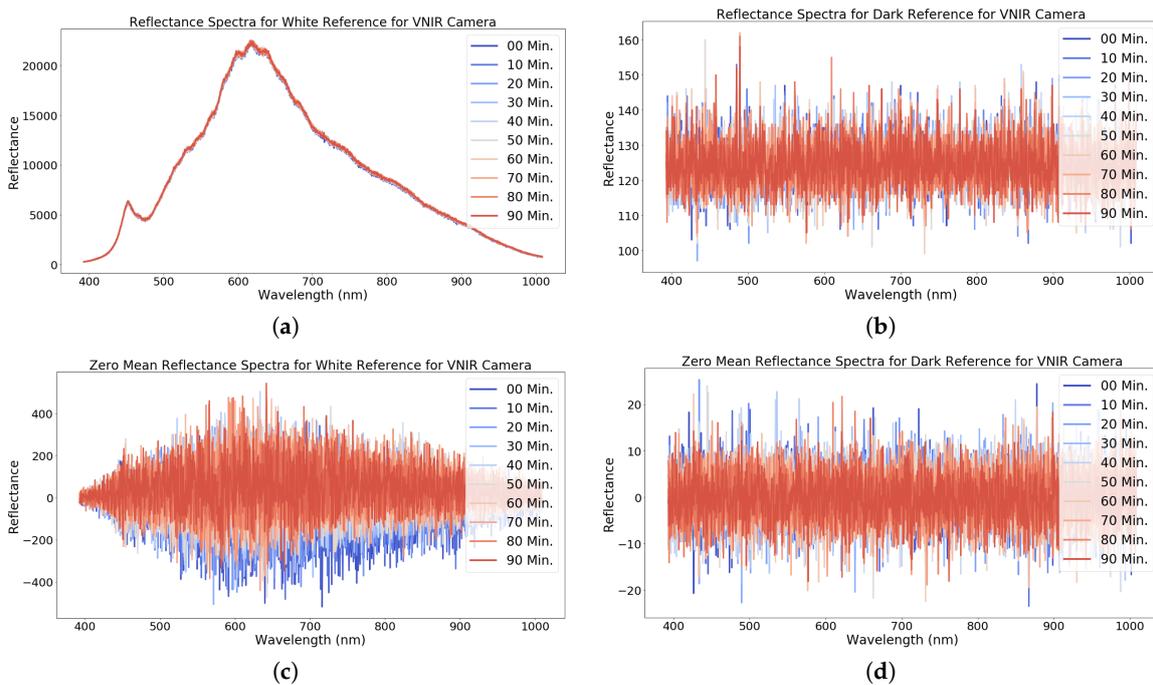


Figure 11. (a) White reference and (b) dark reference variability for a single pixel associated with the Optopolymer reference for a VNIR Camera. Zero mean reflectance for white reference (c) and dark reference (d).

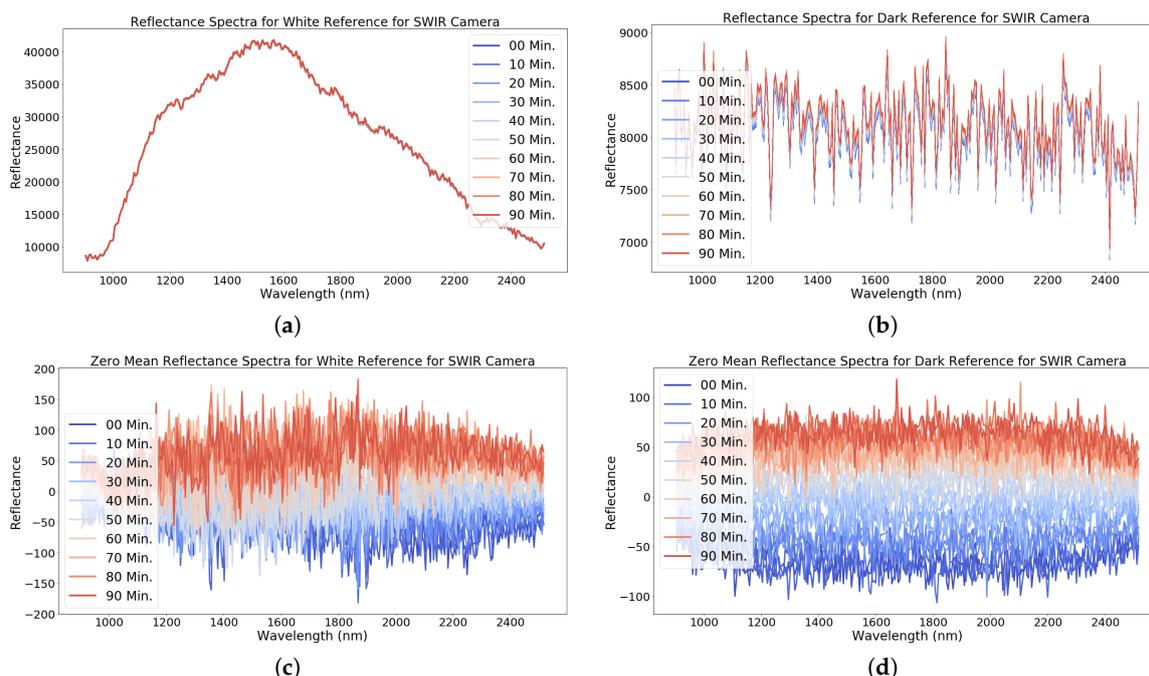


Figure 12. (a) White reference and (b) dark reference variability for a single pixel associated with the Optopolymer reference under controlled conditions for our SWIR Camera. Zero mean reflectance for the white reference (c) and dark reference (d).

The following sections present experiments with real data that demonstrates how using a suitable clustering scheme allows establishing a stochastic hierarchical regression strategy that improves the ability to estimate geometallurgical variables. In this example, samples have been characterized by laboratory tests and hyperspectral images over them have been acquired.

6.2. Geometallurgical Variables Estimation

6.2.1. Database

For this experiment a set of 146 samples, each one coming from 20 m drill core segment composite material has been considered. For each sample a set of variables has been obtained from laboratory metallurgical tests. First 4 variables are taken according to a standard protocol for Rougher flotation, and the last variable from laboratory mills. An extract of the available samples variables is described in Table 1.

Table 1. Measurements for some samples in the available data.

Sample ID	RECCU (%)	RECMO (%)	PH (-)	CONSCAL (Kg/ton)	WI (-)
Sample 001	85.50	63.50	7.70	0.12	14.50
Sample 002	86.6	43.80	8.70	0.16	16.00
Sample 137	90.6	63.3	7.10	0.26	13.00

Measured variables for the available samples correspond to copper (RECCU) and molybdenum (RECMO) mean metallurgical recovery, PH of the sample (PH), calcium carbonate consumption (CONSCAL) and bond working index (WI). These variables are of interest for mine and process planning and are used in the geometallurgical model of the ore deposit. Figure 13 shows histograms for all variables. Note the dynamic range for all of them is expressive enough, allowing all variables suitable for reasonably good modeling.

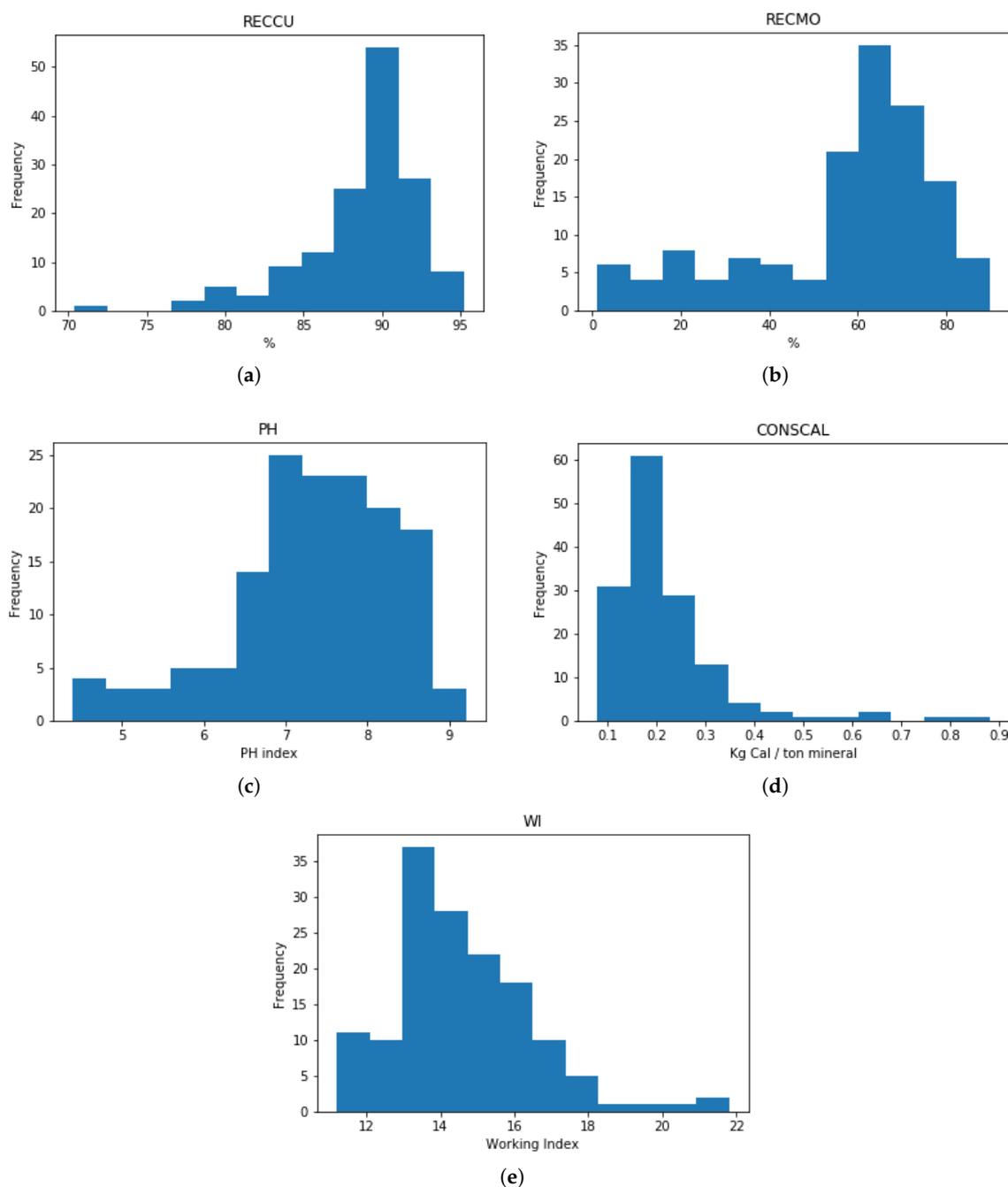


Figure 13. Histograms of the measured variables. Distribution and dynamic ranges for: (a) recovery of copper, (b) recovery of molybdenum, (c) PH of the sample, (d) consumption of calcium carbonate and (e) bond working index.

6.2.2. Variability of Samples in the Spectral Domain Analysis

Each of the available samples correspond to 100 gr. of pulverized material that has been uniformly distributed over a flat surface and a hyperspectral image has been taken using two hyperspectral cameras (VNIR and SWIR). As expected, the result for each sample is a population of spectra that capture both the spectral variability of the content (related with geochemical and mineralogical characteristics of the mineral) and the variability of operating and environmental conditions (see Figure 14). We consider this dynamic is specific to the sample and hence correlating this variability with some external measurement or any associated assay results is possible.

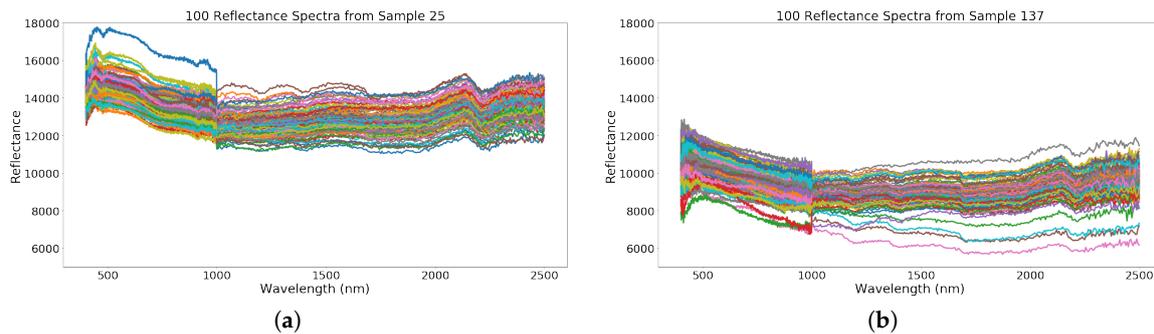


Figure 14. Examples of spectra in the available samples; 100 spectra randomly selected from (a) sample 25 and (b) sample 137. Colormap: random color for each spectrum.

6.2.3. Experiment Scenarios

Results presented in the following section compare the performance of four different scenarios and are intended to test the proposed training and prediction pipelines. The idea is to get some measurements to compare their impact on the quality of the regression models for all five geometallurgical variables.

For the base scenario 1 a naive strategy is presented. It consists of using spectra without dimensionality reduction and without using any clustering method. This approach seeks to establish a machine learning system on the entire database directly, where the original spectra are associated individually to the measured variables.

For the rest of the scenarios different configurations of the proposed methodology are applied. For both the second and the third scenarios, a simple k-means clustering approach and the proposed HySSIC clustering method are used respectively without dimensionality reduction. The fourth scenario finally uses the entire proposed pipeline, i.e., the HySSIC clustering method and the proposed high energy preservation dimensionality reduction method along with the hierarchical regression model. In each case, the HySSIC method is used with the wavelet kernel-trick distance with $a = 1$ to avoid any unnecessary scale change.

6.2.4. Configuration and Results

For the training pipeline, 120 samples were randomly selected from the 146 available samples. For each sample, all five geometallurgical variables are measured. Then 50 patches were taken for each sample for the training process, where each patch corresponds to a small section of 10×10 pixels (100 spectra by patch). Note that each one of these patches shares the same metallurgical measurement (the available measurements for the sample).

The estimation performance of the system was tested on 21 samples that were left out of the training pipeline, applying on them the prediction pipeline. For each sample in this testing data set 30 patches were obtained.

Finally, after applying the prediction pipeline to all patches from the testing sample, a set of histograms showing the experimental distribution of the estimated geometallurgical variables was obtained—Figure 15 shows histograms for sample 137. For this analysis, the mean value of these distributions are considered as the final estimations to be compared to the measured values using root mean square error (RMSE) and mean absolute error (MAE) as performance measures.

Global estimation results for testing data set are summarized in Table 2. Observe how the estimation error is improved for all target variables, compared to naive scenario 1, by adding even a simple clustering strategy to the process—shown in scenario 2. The third scenario results show that an appropriate choice of the clustering technique can also affect results—where HySSIC is slightly better than k-means; this results needs further analysis on other choices for the spectral distance but it is beyond the scope of this work. Scenario 4 shows that the estimation pipeline presents a significant improvement with respect to scenario 1.

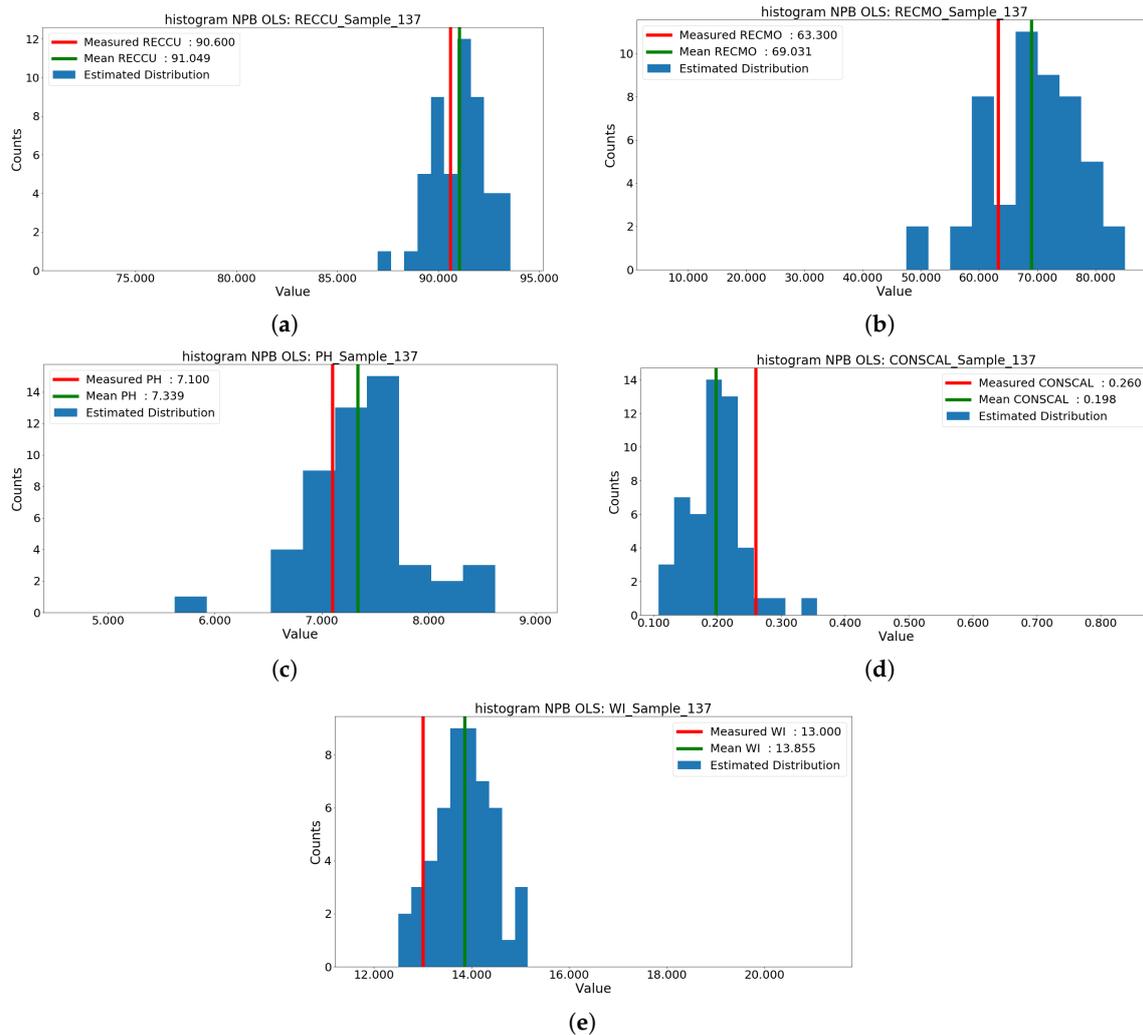


Figure 15. Distribution of the estimated geometallurgical variables for sample 137. Estimations for (a) recovery of copper, (b) recovery of molybdenum, (c) PH, (d) consumption of calcium carbonate and (e) bond working index.

Table 2. Performance of the proposed stochastic hierarchical clustered regression based on hyperspectral images.

Model Regression		RECCU	RECMO	PH	CONSCAL	WI
	Dynamic Range	24.799	85.500	4.799	0.800	10.600
Scenario 1: Naive individual Spectra based classification	MAE	4.568	18.639	0.810	0.095	1.594
	RMSE	5.192	23.387	1.061	0.125	2.077
Scenario 2: Proposed Approach with k-means clustering and no dim. reduction	MAE	1.540	10.224	0.533	0.071	1.385
	RMSE	1.861	12.845	0.751	0.079	1.408
Scenario 3: Proposed Approach with HySSIC Clustering and no dim. reduction	MAE	1.461	8.671	0.351	0.069	0.957
	RMSE	1.651	9.749	0.378	0.075	0.991
Scenario 4: Proposed Approach with HySSIC Clustering and dim. reduction	MAE	0.621	2.398	0.153	0.066	0.350
	RMSE	0.853	3.400	0.189	0.071	0.441

It is interesting to also observe that the estimation error of RECMO variable (Table 2) is slightly worse than the ones for the other variables. This could be related to the fact that the histogram for that estimation (in Figure 15b) can be interpreted as having two modes. Thus, a possible explanation for this error behavior is that there could be two distinguishable populations in the experimental estimation

distribution—besides, in the original variable histogram in Figure 13b a similar bimodal shape can be observed. Therefore, in this case the mean value may not be the best estimator for the RECMO variable. Probably if we chose a more robust estimator from the empirical distribution, as is the case for the median, the result could be improved. On the other hand, this bimodal distribution could be caused by the existence of two different metallurgical phenomena related to the same recovery process. It is precisely this kind of phenomenon that classic spectral analysis can not capture. In this case, the correct estimator should be determined by applying metallurgical expert knowledge, probably with the help of the mineralogical analysis of the samples.

7. Discussion and Future Work

In this work we have presented a survey of several challenges that can be faced when dealing with the characterization of geometallurgical variables using hyperspectral image analysis on both geological and metallurgical samples: spectra inherent variability, spatial and spectral context analysis and especially how to take advantage of the combination and interactions between both contexts to build a robust geometallurgical characterization model.

We have proposed a promising stochastic methodology to characterize geometallurgical variables for metallurgical mineral samples using hyperspectral images, which includes training and prediction pipelines. This methodology relies on three main novel pillars: (a) a preserved energy spectral dimensionality reduction method, (b) a hybrid spatial spectral iterative clustering method and (c) a stochastic hierarchical regression model.

Experiments showed that the proposed preserved energy spectral dimensionality reduction method improves not only the computational processing time—and the computer memory workload—but also improves the whole pipeline performance. This shows the method is able to clean up spurious information contained in individual spectra, preventing the method from overfitting noisy spectra or some other artifact coming from image capture field conditions.

The proposed clustering method showed the helpful ability to combine spatial and spectral information, to take advantage of structural information differences in metallurgical samples, such as grain profile and target phenomenon sparsity. Experiments showed that an appropriate choice of the clustering method in the proposed methodology has a remarkable impact on the output performance for the pipeline. On the other hand, it worth mentioning that the so called kernel trick used in this method may provide it with a brand new field for researching topological properties, when assuming that spectra can be in more complex theoretical spaces. The latter, for instance, can be an interesting and attractive way for finding new spectral features for known elements.

Finally, the stochastic hierarchical regression model provides the methodology with the ability to characterize the inherent variability of spectra captured in field conditions. In this sense, experiments showed that adding an additional level to the regression model improves significantly the prediction pipeline performance, indicating that spectral distributions, rather than isolated ones, are a very advisable choice to predict geometallurgical variables.

Author Contributions: Conceptualization, A.F.E., A.E. and F.A.S.-L.; methodology, A.F.E., A.E. and F.A.S.-L.; software, F.A.S.-L., C.V. and S.L.; validation, C.V., F.A.S.-L. and S.L.; formal analysis, A.F.E., F.A.S.-L., A.E. and S.L.; investigation, A.F.E., A.E., F.A.S.-L., G.D. and S.L.; resources, C.V., S.L. and F.A.S.-L.; data curation, C.V., F.A.S.-L.; writing—original draft preparation, F.A.S.-L., S.L., G.D. and C.V.; writing—review and editing, A.F.E., A.E. and F.A.S.-L.; visualization, C.V. and F.S.; supervision, A.F.E. and A.E.; project administration, A.F.E. and A.E. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported and partially funded by The Advanced Mining Technology Center (AMTC) Basal project (ANID/PIA Project AFB180004) and CONICYT PHD Fellowship 2013 (21130890).

Acknowledgments: The authors are grateful to the contribution from the Outlier SpA Company; Cristian Jara for supporting the experimental analysis; and the ALGES laboratory team for their fruitful and ever-present support.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ALGES	Advanced Laboratory for Geostatistical Supercomputing
AMTC	Advanced Mining Technology Center
CONSCAL	calcium carbonate consumption
HySSIC	hybrid spectral spatial iterative clustering
HSI	hyperspectral image
LMM	linear mixed model
MAE	mean absolute error
MSI	multispectral image
NIR	near infrared
RGB	red-green-blue
RECCU	recovered soluble copper
RECMO	molybdenum recovery
RMSE	root mean square error
SLIC	simple linear iterative clustering
SWIR	short wave infra red
VNIR	visible and near infrared
WI	bond working index

References

- van der Meer, F. Near-infrared laboratory spectroscopy of mineral chemistry: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *65*, 71–78. [CrossRef]
- Popelka-Filcoff, R.S.; Mauger, A.; Lenehan, C.E.; Walshe, K.; Pring, A. HyLogger™ near-infrared spectral analysis: A non-destructive mineral analysis of Aboriginal Australian objects. *Anal. Methods* **2014**, *6*, 1309–1316. [CrossRef]
- Martini, B.A.; Bellian, J.; Katz, D.; Fonteneau, L.C.; Carey, R.; Guisinger, M.; Nordeng, S.H. Continuous mineralogical characterization of the bakken-three forks formations: New geological insights from hyperspectral core imaging. In Proceedings of the SPE/AAPG/SEG Unconventional Resources Technology Conference 2020, URTEC 2020, Denver, CO, USA, 22–24 July 2019. [CrossRef]
- Martini, B.A.; Harris, A.C.; Carey, R.; Goodey, N.; Honey, F.; Tuffilli, N. Automated Hyperspectral Core Imaging-A Revolutionary New Tool for Exploration, Mining and Research. *Proc. Explor.* **2017**, *17*, 911–922.
- Plaue, J.W.; Klunder, G.L.; Hutcheon, I.D.; Czerwinski, K.R. Near infrared reflectance spectroscopy as a process signature in uranium oxides. *J. Radioanal. Nucl. Chem.* **2013**, *296*, 551–555. [CrossRef]
- Analytical Spectral Devices, Inc. *ASD Technical Guide*, 3rd ed.; ASD Inc.: Boulder, CO, USA, 1999.
- Malvern Panalytical. ASD LabSpec: NIRS Spectrometers and Spectrophotometers. 2020. Available online: <https://www.malvernpanalytical.com/en/products/product-range/asd-range/labspec-range/> (accessed on 4 November 2020).
- Berman, M.; Kiiveri, H.; Lagerstrom, R.; Ernst, A.; Dunne, R.; Huntington, J.F. ICE: A statistical approach to identifying endmembers in hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2085–2095. [CrossRef]
- Keshava, N.; Mustard, J.F. Spectral unmixing. *Signal Process. Mag. IEEE* **2002**, *19*, 44–57.
- Bian, X.; Li, S.; Shao, X.; Liu, P. Variable space boosting partial least squares for multivariate calibration of near-infrared spectroscopy. *Chemom. Intell. Lab. Syst.* **2016**, *158*, 174–179. [CrossRef]
- Karaca, A.C.; Erturk, A.; Gullu, M.; Elmas, M.; Erturk, S. Plastic waste sorting using infrared hyperspectral imaging system. In Proceedings of the 2013 21st Signal Processing and Communications Applications Conference (SIU 2013), Haspolat, Turkey, 24–26 April 2013; pp. 1–4.
- Serranti, S.; Gargiulo, A.; Bonifazi, G. Characterization of post-consumer polyolefin wastes by hyperspectral imaging for quality control in recycling processes. *Waste Manag.* **2011**, *31*, 2217–2227.
- Serranti, S.; Gargiulo, A.; Bonifazi, G. The utilization of hyperspectral imaging for impurities detection in secondary plastics. *Open Waste Manag. J.* **2010**, *3*, 56–70.

14. Elmasry, G.; Kamruzzaman, M.; Sun, D.W.; Allen, P. Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: A review. *Crit. Rev. Food Sci. Nutr.* **2012**, *52*, 999–1023.
15. Sun, D.W. *Hyperspectral Imaging for Food Quality Analysis and Control*; Elsevier: Cambridge, MA, USA, 2010.
16. Gowen, A.; O'Donnell, C.; Cullen, P.; Downey, G.; Frias, J. Hyperspectral imaging—An emerging process analytical tool for food quality and safety control. *Trends Food Sci. Technol.* **2007**, *18*, 590–598.
17. Boggs, J.L.; Tsegaye, T.; Coleman, T.L.; Reddy, K.; Fahsi, A. Relationship between hyperspectral reflectance, soil nitrate-nitrogen, cotton leaf chlorophyll, and cotton yield: A step toward precision agriculture. *J. Sustain. Agric.* **2003**, *22*, 5–16.
18. Xu, P.; Liu, J.; Xue, L.; Zhang, J.; Qiu, B. Adaptive Grouping Distributed Compressive Sensing Reconstruction of Plant Hyperspectral Data. *Sensors* **2017**, *17*, 1322. [[CrossRef](#)]
19. Berné, O.; Helens, A.; Pilleri, P.; Joblin, C. Non-negative matrix factorization pansharpening of hyperspectral data: An application to mid-infrared astronomy. In Proceedings of the 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.
20. Hege, E.K.; O'Connell, D.; Johnson, W.; Basty, S.; Dereniak, E.L. Hyperspectral imaging for astronomy and space surveillance. In *Optical Science and Technology, SPIE's 48th Annual Meeting*; International Society for Optics and Photonics: Bellingham, WA, USA, 2004; pp. 380–391.
21. Lillesand, T.M.; Kiefer, R.W.; Chipman, J.W. *Remote Sensing and Image Interpretation*, 7th ed.; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 2015; Chapter 8.
22. Melgani, F.; Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *Geosci. Remote Sens. IEEE Trans.* **2004**, *42*, 1778–1790.
23. Brando, V.E.; Dekker, A.G. Satellite hyperspectral remote sensing for estimating estuarine and coastal water quality. *Geosci. Remote Sens. IEEE Trans.* **2003**, *41*, 1378–1387.
24. Malvern Panalytical. ASD Terraspec: Portable and Handheld Mineral NIR Analyzers. 2020. Available online: <https://www.malvernpanalytical.com/en/products/product-range/asd-range/terraspec-range> (accessed on 15 November 2020).
25. El Sobky, M.A.; Madani, A.A.; Surour, A.A. Spectral characterization of the Batuga granite pluton, South Eastern Desert, Egypt: Influence of lithological and mineralogical variation on ASD Terraspec data. *Arab. J. Geosci.* **2020**, *13*. [[CrossRef](#)]
26. Zhou, X.; Jara, C.; Bardoux, M.; Plasencia, C. Multi-Scale integrated application of Spectral Geology and Remote Sensing for Mineral Exploration. *Proc. Explor.* **2017**, *17*, 899–910.
27. Specim. Hyperspectral Imaging Cameras and Systems. 2020. Available online: <https://www.specim.fi> (accessed on 14 November 2020).
28. Horstrand, P.; Guerra, R.; Rodríguez, A.; Díaz, M.; López, S.; López, J.F. A UAV Platform Based on a Hyperspectral Sensor for Image Capturing and On-Board Processing. *IEEE Access* **2019**, *7*, 66919–66938. [[CrossRef](#)]
29. Behmann, J.; Acebron, K.; Emin, D.; Bennertz, S.; Matsubara, S.; Thomas, S.; Bohnenkamp, D.; Kuska, M.T.; Jussila, J.; Salo, H.; et al. Specim IQ: Evaluation of a New, Miniaturized Handheld Hyperspectral Camera and Its Application for Plant Phenotyping and Disease Detection. *Sensors* **2018**, *18*, 441. [[CrossRef](#)]
30. Cilia, C.; Panigada, C.; Rossini, M.; Meroni, M.; Busetto, L.; Amaducci, S.; Boschetti, M.; Picchi, V.; Colombo, R. Nitrogen Status Assessment for Variable Rate Fertilization in Maize through Hyperspectral Imagery. *Remote Sens.* **2014**, *6*, 6549–6565. [[CrossRef](#)]
31. Kuula, J.; Pölönen, I.; Puupponen, H.H.; Selander, T.; Reinikainen, T.; Kalenius, T.; Saari, H. Using VIS/NIR and IR spectral cameras for detecting and separating crime scene details. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XI*; Carapezza, E.M., Ed.; International Society for Optics and Photonics: Bellingham, WA, USA, 2012; Volume 8359, pp. 83590P-1–83590P-11. [[CrossRef](#)]
32. Angel, Y.; Turner, D.; Parkes, S.; Malbeteau, Y.; Lucieer, A.; McCabe, M.F. Automated Georectification and Mosaicking of UAV-Based Hyperspectral Imagery from Push-Broom Sensors. *Remote Sens.* **2019**, *12*, 34. [[CrossRef](#)]
33. Turner, D.; Lucieer, A.; McCabe, M.; Parkes, S.; Clarke, I. Pushbroom hyperspectral imaging from an unmanned aircraft system (UAS)—Geometric processing workflow and accuracy assessment. *ISPRS Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-2/W6*, 379–384. [[CrossRef](#)]

34. Xue, L. Application of IDL and ENVI Redevelopment in Hyperspectral Image Preprocessing. In *Computer and Computing Technologies in Agriculture IV*; Li, D.; Liu, Y.; Chen, Y., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 403–409.
35. Neal, L.C.; Wilkinson, J.J.; Mason, P.J.; Chang, Z. Spectral characteristics of propylitic alteration minerals as a vectoring tool for porphyry copper deposits. *J. Geochem. Explor.* **2018**, *184*, 179–198. [[CrossRef](#)]
36. Percival, J.B.; Bosman, S.A.; Potter, E.G.; Peter, J.M.; Laudadio, A.B.; Abraham, A.C.; Shiley, D.A.; Sherry, C. Customized spectral libraries for effective mineral exploration: Mining national mineral collections. *Clays Clay Miner.* **2018**, *66*, 297–314. [[CrossRef](#)]
37. Schneider, S.; Murphy, R.J.; Melkumyan, A. Evaluating the performance of a new classifier—The GP-OAD: A comparison with existing methods for classifying rock type and mineralogy from hyperspectral imagery. *ISPRS J. Photogramm. Remote Sens.* **2014**, *98*, 145–156. [[CrossRef](#)]
38. Tarabalka, Y.; Chanussot, J.; Benediktsson, J. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognit.* **2010**, *43*, 2367–2379. [[CrossRef](#)]
39. Lishchuk, V.; Koch, P.H.; Ghorbani, Y.; Butcher, A.R. Towards integrated geometallurgical approach: Critical review of current practices and future trends. *Miner. Eng.* **2020**, *145*, 106072. [[CrossRef](#)]
40. Bergh, L.G.; Yianatos, J.B. The long way toward multivariate predictive control of flotation processes. *J. Process Control* **2011**, *21*, 226–234. [[CrossRef](#)]
41. Perez, C.A.; Estévez, P.A.; Vera, P.A.; Castillo, L.E.; Aravena, C.M.; Schulz, D.A.; Medina, L.E. Ore grade estimation by feature selection and voting using boundary detection in digital image analysis. *Int. J. Miner. Process.* **2011**, *101*, 28–36. [[CrossRef](#)]
42. Howarth, D.F.; Rowlands, J.C. Quantitative assessment of rock texture and correlation with drillability and strength properties. *Rock Mech. Rock Eng.* **1987**, *20*, 57–85. [[CrossRef](#)]
43. Yu, H.; Zhang, X.; Wang, S.; Hou, B. Context-Based Hierarchical Unequal Merging for SAR Image Segmentation. *Geosci. Remote Sens. IEEE Trans.* **2013**, *51*, 995–1009. [[CrossRef](#)]
44. Giraud, R.; Berthoumieu, Y. Texture Superpixel Clustering from Patch-based Nearest Neighbor Matching. In Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2–6 September 2019; pp. 1–5. [[CrossRef](#)]
45. Mirmehdi, M. *Handbook of Texture Analysis*; Imperial College Press: London, UK, 2008.
46. Haggerty, S.E. The ore minerals and their intergrowths. *Geochim. Cosmochim. Acta* **1983**, *47*, 992–993. [[CrossRef](#)]
47. Hunt, G.R.; Vincent, R.K. The behavior of spectral features in the infrared emission from particulate surfaces of various grain sizes. *J. Geophys. Res.* **1968**, *73*, 6039–6046. [[CrossRef](#)]
48. Myers, T.L.; Brauer, C.S.; Su, Y.F.; Blake, T.A.; Tonkyn, R.G.; Ertel, A.B.; Johnson, T.J.; Richardson, R.L. Quantitative reflectance spectra of solid powders as a function of particle size. *Appl. Opt.* **2015**, *54*, 4863. [[CrossRef](#)] [[PubMed](#)]
49. Le Bras, A.; Erard, S. Reflectance spectra of regolith analogs in the mid-infrared: Effects of grain size. *Planet. Space Sci.* **2003**, *51*, 281–294. [[CrossRef](#)]
50. Cooper, B.L.; Salisbury, J.W.; Killen, R.M.; Potter, A.E. Midinfrared spectral features of rocks and their powders. *J. Geophys. Res. E Planets* **2002**, *107*. [[CrossRef](#)]
51. Xie, F.; Fan, Y.; Zhou, M. Dimensionality reduction by weighted connections between neighborhoods. *Abstr. Appl. Anal.* **2014**, *2014*. [[CrossRef](#)]
52. Nair, P.; Chaudhury, K.N. Fast High-Dimensional Bilateral and Nonlocal Means Filtering. *IEEE Trans. Image Process.* **2019**, *28*, 1470–1481. [[CrossRef](#)]
53. Armi, L.; Fekri-Ershad, S. Texture image analysis and texture classification methods—A review. *arXiv* **2019**, arXiv:1904.06554.
54. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [[CrossRef](#)]
55. Chang, T.; Kuo, C.C. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Image Process.* **1993**, *2*, 429–441. [[CrossRef](#)] [[PubMed](#)]
56. Jain, A.K.; Farrokhnia, F. Unsupervised texture segmentation using Gabor filters. *Pattern Recognit.* **1991**, *24*, 1167–1186. [[CrossRef](#)]
57. Lobos, R.; Silva, J.F.; Ortiz, J.M.; Díaz, G.; Egaña, A. Analysis and classification of natural rock textures based on new transform-based features. *Math. Geosci.* **2016**, *48*, 835–870. [[CrossRef](#)]

58. Diaz, G.F.; Ortiz, J.M.; Silva, J.F.; Lobos, R.A.; Egana, A.F. Variogram-Based Descriptors for Comparison and Classification of Rock Texture Images. *Math. Geosci.* **2020**, *52*, 451–476. [[CrossRef](#)]
59. Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [[CrossRef](#)]
60. Do, M.N.; Vetterli, M. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. *IEEE Trans. Image Process.* **2002**, *11*, 146–158. [[CrossRef](#)]
61. Varanasi, M.K.; Aazhang, B. Parametric generalized Gaussian density estimation. *J. Acoust. Soc. Am.* **1989**, *86*, 1404–1415. [[CrossRef](#)]
62. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [[CrossRef](#)]
63. Szolgay, D.; Szirányi, T. Adaptive image decomposition into cartoon and texture parts optimized by the orthogonality criterion. *IEEE Trans. Image Process.* **2012**, *21*, 3405–3415. [[CrossRef](#)]
64. Yi, S.; Labate, D.; Easley, G.R.; Krim, H. A shearlet approach to edge analysis and detection. *IEEE Trans. Image Process.* **2009**, *18*, 929–941.
65. Kutyniok, G.; Labate, D. *Shearlets: Multiscale Analysis for Multivariate Data*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
66. Chiles, J.P.; Delfiner, P. *Geostatistics: Modeling Spatial Uncertainty*; John Wiley & Sons: Hoboken, NJ, USA, 2009; Volume 497.
67. Ziegel, E.R.; Deutsch, C.; Journel, A. GSLIB: Geostatistical Software Library and User's Guide. *Technometrics* **1995**, *37*, 126. [[CrossRef](#)]
68. Egaña, A.F.; Ortiz, J.M. Image Segmentation for Mineral Identification in an Oxide Copper Deposit abstract. In Proceedings of the 4th International Conference on Mining Innovation, MININ 2010, Santiago, Chile, 23–25 June 2010; pp. 405–412.
69. Delon, J.; Desolneux, A.; Lisani, J.L.; Petro, A.B. Color image segmentation using acceptable histogram segmentation. *Lect. Notes Comput. Sci.* **2005**, *3523*, 239–246. [[CrossRef](#)]
70. Delon, J.; Desolneux, A.; Lisani, J.L.; Petro, A.B. A Nonparametric Approach for Histogram Segmentation. *IEEE Trans. Image Process.* **2007**, *16*, 253–261. [[CrossRef](#)] [[PubMed](#)]
71. Broomhead, D.S.; Kirby, M. A New Approach to Dimensionality Reduction: Theory and Algorithms. *SIAM J. Appl. Math.* **2000**, *60*, 2114–2142. [[CrossRef](#)]
72. Bi, Y.; Xie, Q.; Peng, S.; Tang, L.; Hu, Y.; Tan, J.; Zhao, Y.; Li, C. Dual stacked partial least squares for analysis of near-infrared spectra. *Anal. Chim. Acta* **2013**, *792*, 19–27. [[CrossRef](#)]
73. Boulesteix, A.L. PLS Dimension Reduction for Classification with Microarray Data. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 1–30. [[CrossRef](#)]
74. Wen, T.; Zhang, Z. Deep Convolution Neural Network and Autoencoders-Based Unsupervised Feature Learning of EEG Signals. *IEEE Access* **2018**, *6*, 25399–25410. [[CrossRef](#)]
75. Wang, Y.; Yao, H.; Zhao, S. Auto-encoder based dimensionality reduction. *Neurocomputing* **2016**, *184*, 232–242. [[CrossRef](#)]
76. San Martin, G.; López Droguett, E.; Meruane, V.; das Chagas Moura, M. Deep variational auto-encoders: A promising tool for dimensionality reduction and ball bearing elements fault diagnosis. *Struct. Health Monit.* **2019**, *18*, 1092–1128. [[CrossRef](#)]
77. Daubechies, I. Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* **1988**, *41*, 909–996. [[CrossRef](#)]
78. Daubechies, I. Orthonormal Bases of Compactly Supported Wavelets II. Variations on a Theme. *SIAM J. Math. Anal.* **1993**, *24*, 499–519. [[CrossRef](#)]
79. Chen, Z.; Liu, J.; Deng, Y.; He, K.; Hopcroft, J.E. Adaptive Wavelet Clustering for Highly Noisy Data. In Proceedings of the 2019 IEEE 35th International Conference on Data Engineering (ICDE), Macao, China, 8–12 April 2019; pp. 328–337. [[CrossRef](#)]
80. Qu, Y.; Adam, B.L.; Thornquist, M.; Potter, J.D.; Thompson, M.L.; Yasui, Y.; Davis, J.; Schellhammer, P.F.; Cazares, L.; Clements, M.A.; et al. Data reduction using a discrete wavelet transform in discriminant analysis of very high dimensionality data. *Biometrics* **2003**, *59*, 143–151. [[CrossRef](#)] [[PubMed](#)]
81. Bruce, L.M.; Koger, C.H.; Jiang Li. Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2331–2338. [[CrossRef](#)]

82. Wickmann, J. A Wavelet Approach to Dimension Reduction and Classification Of Hyperspectral Data. Master's Thesis, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway, 2007; pp. 1–167.
83. Kriegel, H.P.; Kröger, P.; Sander, J.; Zimek, A. Density-based clustering. *WIREs Data Min. Knowl. Discov.* **2011**, *1*, 231–240. [[CrossRef](#)]
84. Murty, M.N.; Flynn, P.J.; Jain, A.K. Data Clustering: A Review. *ACM Comput. Surv.* **1999**, *31*, 43.
85. Hai, M.; Zhang, Y.; Li, H. A Performance Comparison of Big Data Processing Platform Based on Parallel Clustering Algorithms. *Procedia Comput. Sci.* **2018**, *139*, 127–135. [[CrossRef](#)]
86. Rezaei, M. Improving a Centroid-Based Clustering by Using Suitable Centroids from Another Clustering. *J. Classif.* **2020**, *37*, 352–365. [[CrossRef](#)]
87. Zhengqin Li.; Jiansheng Chen. Superpixel segmentation using Linear Spectral Clustering. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1356–1363. [[CrossRef](#)]
88. Chen, J.; Li, Z.; Huang, B. Linear Spectral Clustering Superpixel. *IEEE Trans. Image Process.* **2017**, *26*, 3317–3330. [[CrossRef](#)]
89. Swarndepp, S.J.; Pandya, S. An Overview of Partitioning Algorithms in Clustering Techniques. *Int. J. Adv. Res. Comput. Technol. (IJARCET)* **2016**, *5*, 1943–1946.
90. Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J. Spectral–spatial classification of hyperspectral imagery based on partitioning clustering techniques. *Geosci. Remote Sens. IEEE Trans.* **2009**, *47*, 2973–2987. [[CrossRef](#)]
91. Tarabalka, Y.; Chanussot, J.; Benediktsson, J.A. Segmentation and Classification of Hyperspectral Images Using Minimum Spanning Forest Grown From Automatically Selected Markers. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2010**, *40*, 1267–1279. [[CrossRef](#)] [[PubMed](#)]
92. Taskesen, B.; Koz, A.; Alatan, A.; Weatherbee, O. Change Detection for Hyperspectral Images Using Extended Mutual Information and Oversegmentation. In Proceedings of the 2018 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, The Netherlands, 23–26 September 2018; pp. 1–5. [[CrossRef](#)]
93. Wang, H.; Peng, X.; Xiao, X.; Liu, Y. BSLIC: SLIC Superpixels Based on Boundary Term. *Symmetry* **2017**, *9*, 31. [[CrossRef](#)]
94. Stutz, D. Superpixel Segmentation: An Evaluation. *Lect. Notes Comput. Sci. (Lect. Notes Artif. Intell. Lect. Notes Bioinform.)* **2015**, *9358*, 555–562. [[CrossRef](#)]
95. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Susstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *Pattern Anal. Mach. Intell. IEEE Trans.* **2012**, *34*, 2274–2282. [[CrossRef](#)]
96. Tilton, J.; Pasolli, E. Incorporating edge information into best merge region-growing segmentation. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014; pp. 4891–4894. [[CrossRef](#)]
97. Rodriguez, M.; Comin, C.; Casanova, D.; Bruno, O.; Amancio, D.; da Costa, L.; Rodrigues, F. Clustering algorithms: A comparative approach. *PLoS ONE* **2019**, *14*, e0210236. [[CrossRef](#)]
98. Bicego, M.; Cristani, M.; Fusiello, A.; Murino, V. Watershed-based unsupervised clustering. *Lect. Notes Comput. Sci. (Lect. Artif. Intell. Lect. Notes Bioinform.)* **2003**, *2683*, 83–94. [[CrossRef](#)]
99. Theiler, J.P.; Gisler, G. Contiguity-enhanced k-means clustering algorithm for unsupervised multispectral image segmentation. In *Optical Science, Engineering and Instrumentation'97*; International Society for Optics and Photonics: 1997; pp. 108–118.
100. Tsapanos, N.; Tefas, A.; Nikolaidis, N.; Pitas, I. Kernel matrix trimming for improved Kernel K-means clustering. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 2285–2289.
101. Haroske, D.D.; Skandera, P.; Triebel, H. An Approach to Wavelet Isomorphisms of Function Spaces via Atomic Representations. *J. Fourier Anal. Appl.* **2018**, *24*, 830–871. [[CrossRef](#)]
102. Krupnik, D.; Khan, S. Close-range, ground-based hyperspectral imaging for mining applications at various scales: Review and case studies. *Earth-Sci. Rev.* **2019**, *198*, 102952. [[CrossRef](#)]
103. Clark, R.N.; Roush, T.L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.* **1984**, *89*, 6329–6340. [[CrossRef](#)]

104. Clark, R.N.; King, T.V.; Gorelick, N.S. Automatic Continuum Analysis of Reflectance Spectra. In Proceedings of the 3rd Airborne Imaging Spectrometer Data Analysis Workshop (JPL), NASA JPL Publication, 1987; pp. 138–142.
105. Kruse, F.A.; Raines, G.L.; Watson, K. Analytical techniques for extracting geologic information from multichannel airborne spectroradiometer and imaging spectrometer data. In Proceedings of the International Symposium on Remote Sensing of Environment, Fourth Thematic Conference: Remote Sensing for Exploration Geology, San Francisco, CA, USA, 1–4 April 1986; pp. 309–324.
106. Green, A.A.; Craig, M.D. Analysis of aircraft spectrometer data with logarithmic residuals. In Proceedings of the Airborne Imaging Spectrometer Data Analysis Workshop, Pasadena, CA, USA, 8–10 April 1985; pp. 111–119.
107. Yousefi, B.; Sojasi, S.; Castanedo, C.I.; Maldague, X.P.; Beaudoin, G.; Chamberland, M. Comparison assessment of low rank sparse-PCA based-clustering/classification for automatic mineral identification in long wave infrared hyperspectral imagery. *Infrared Phys. Technol.* **2018**, *93*, 103–111. [[CrossRef](#)]
108. Pisapia, C.; Jamme, F.; Duponchel, L.; Ménez, B. Tracking hidden organic carbon in rocks using chemometrics and hyperspectral imaging. *Sci. Rep.* **2018**, *8*, 2396. [[CrossRef](#)] [[PubMed](#)]
109. Ekanayake, E.; Vithana, S.; Ekanayake, E.; Rathnayake, A.; Abeysekara, A.; Oorloff, T.; Herath, H.; Godaliyadda, G.; Ekanayake, M.; Senaratne, A. Mapping ilmenite deposit in Pulmudai, Sri Lanka using a hyperspectral imaging-based surface mineral mapping method. *J. Natl. Sci. Found. Sri Lanka* **2019**, *47*, 271. [[CrossRef](#)]
110. Yousefi, B.; Castanedo, C.I.; Maldague, X.P.; Beaudoin, G. Assessing the reliability of an automated system for mineral identification using LWIR Hyperspectral Infrared imagery. *Miner. Eng.* **2020**, *155*, 106409. [[CrossRef](#)]
111. Johnstone, I.M.; Titterton, D.M. Statistical challenges of high-dimensional data. *Philos. Trans. R. Soc. A Math. Eng. Sci.* **2009**, *367*, 4237–4253. [[CrossRef](#)] [[PubMed](#)]
112. Fan, J.; Li, R. Statistical challenges with high dimensionality: Feature selection in knowledge discovery. In Proceedings of the International Congress of Mathematicians, Madrid, Spain, 22–30 August 2006; European Mathematical Society Publishing House: Zuerich, Switzerland, 2006; Volume 3, pp. 595–622. [[CrossRef](#)]
113. Paoletti, M.; Haut, J.; Plaza, J.; Plaza, A. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 279–317. [[CrossRef](#)]
114. Ma, W.; Gong, C.; Hu, Y.; Meng, P.; Xu, F. The Hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around Qinghai Lake. In *International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Spectrometer Technologies and Applications*; Zhang, L.; Yang, J., Eds.; International Society for Optics and Photonics: Bellingham, WA, USA, 2013; Volume 8910, p. 89101G. [[CrossRef](#)]
115. Jing Wang; Chein-I Chang. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1586–1600. [[CrossRef](#)]
116. Candès, E.J.; Romberg, J.; Tao, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **2006**, *52*, 489–509. [[CrossRef](#)]
117. Hill, B.; Roger, T.; Vorhagen, F.W. Comparative Analysis of the Quantization of Color Spaces on the Basis of the CIELAB Color-Difference Formula. *ACM Trans. Graph.* **1997**, *16*, 109–154. [[CrossRef](#)]
118. Apsemidis, A.; Psarakis, S.; Moguerza, J.M. A review of machine learning kernel methods in statistical process monitoring. *Comput. Ind. Eng.* **2020**, *142*, 106376. [[CrossRef](#)]
119. Pillonetto, G.; Dinuzzo, F.; Chen, T.; De Nicolao, G.; Ljung, L. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica* **2014**, *50*, 657–682. [[CrossRef](#)]
120. Schölkopf, B. The Kernel Trick for Distances. In *Advances in Neural Information Processing Systems 13*; Leen, T.K.; Dietterich, T.G.; Tresp, V., Eds.; MIT Press: Cambridge, MA, USA, 2001; pp. 301–307.
121. Martín-Merino, M.; Muñoz, A. Extending the SOM algorithm to non-Euclidean distances via the kernel trick. *Lect. Notes Comput. Sci. (Lect. Artif. Intell. Lect. Notes Bioinform.)* **2004**, *3316*, 150–157. [[CrossRef](#)]
122. Badshah, N.; Ahmad, A.; Rehman, F. Variational level set image segmentation model coupled with kernel distance function. *J. Algorithms Comput. Technol.* **2020**, *14*. [[CrossRef](#)]
123. Lindstrom, M.J.; Bates, D.M. Newton—Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data. *J. Am. Stat. Assoc.* **1988**, *83*, 1014–1022. [[CrossRef](#)]

124. Bates, D.M.; DebRoy, S. Linear mixed models and penalized least squares. *J. Multivar. Anal.* **2004**, *91*, 1–17. [[CrossRef](#)]
125. Molenberghs, G.; Renard, D.; Verbeke, G. A review of generalized linear mixed models. *J. La Soc. Française Stat.* **2002**, *143*, 53–78.
126. Lowekamp, B.C.; Chen, D.T.; Ibáñez, L.; Blezek, D. The Design of SimpleITK. *Front. Neuroinform.* **2013**, *7*. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).