Improved deconvolution of mineral reflectance spectra

Ronan Rialland, Charles Soussen, Rodolphe Marion, Véronique Carrère

Improved deconvolution of mineral reflectance spectra

Ronan Rialland, Charles Soussen, *Member, IEEE,* Rodolphe Marion, and Véronique Carrère

***Abstract*—The identification and characterization of a mineral from its reflectance spectrum is based on the analysis of the continuum (*i.e.*, the overall shape) and its diagnostic absorptions. The development of automatic methods to retrieve such useful information is a key challenge as new airborne and satellite-borne sensors are developed. We present a flexible and automatic deconvolution procedure able to deal with various minerals and sensors. The approach is based on a physical model and offers the possibility to include noise statistics. It is composed of three successive steps: *(i)* the continuum is pre-estimated and removed using a non-linear least-squares algorithm with adapted constraints; *(ii)* the absorptions are pre-estimated using a greedy algorithm; *(iii)* the continuum and absorption estimates are jointly refined using a non-linear least-squares procedure. Each step is validated first on synthetic spectra, including the sensitivity of the full procedure to instrumental noise. Then, the procedure is applied to diverse laboratory spectra. In most cases, absorption positions are recovered with an accuracy lower than 5 nm, enabling mineral identification. The potential of the proposed deconvolution procedure is finally assessed on hyperspectral images acquired over quarries during a dedicated airborne campaign. Here, minerals such as calcite and gypsum are clearly identified based on their diagnostic absorption features, including when they are in a mixture. Small changes in the shape of the kaolinite doublet are also clearly detected and could be related to crystallinity or mixture with other minerals such as gibbsite.**

***Index Terms*—Mineral reflectance spectra, Hyperspectral images, HySpex, EGO model, AGM procedure.**

I. Introduction

A. Spectroscopy, mineralogy and hyperspectral imaging

HYPERSPECTRAL imaging in the solar reflective domain (VNIR, Visible Near-InfraRead $[400-1300]$ nm and SWIR, Short-Wave InfraRed $[1300-2500]$ nm) provides a reflectance spectrum for each pixel of an image allowing one to retrieve and map physico-chemical properties of the observed surface. In particular, spectra of minerals are composed of specific or diagnostic absorption bands, the position of which mainly depends on the chemical composition [1]. Absorptions in the SWIR are generally narrow and deep, whereas in the VNIR they tend to be broader and weaker because they correspond to various physical phenomena: rotation, vibration, stretching and bending of molecule bonds in the SWIR, crystal field, electronic transitions in the VNIR. The overall shape of a mineral reflectance spectrum, called continuum, depends on surface conditions (*e.g*, grain size, roughness, humidity) and illumination. The shape and strength of the diagnostic absorptions depend on concentrations, potential mixtures and observation conditions, so that a simple comparison of a spectrum with an extensive database using an arbitrary metric generally fails. In the following, we focus on a set of minerals of interest [2] that can be identified and characterized from their reflectance spectra and are often encountered in airborne and satellite-borne hyperspectral images.

New airborne and satellite-borne sensors such as the Airborne Visible InfraRed Imaging Spectrometer, Classic (AVIRIS) [3] and Next Generation (AVIRIS-NG) [4], the HySpex sensors such as Mjölnir (https://www.hyspex.com), the *PRecursore IperSpettrale della Missione Applicativa* (PRISMA) [5] or the Environmental Mapping and Analysis Program (EnMAP) [6] gave rise to an increased number of hyperspectral images over large areas of mineralogical interest, so that detailed spectral analysis can be performed. Developing flexible and automatic algorithms able to deal with various sensor characteristics and a wide variety of minerals is an important objective which still remains challenging.

B. Models and methods

Several approaches (see [7] and references therein for a review) have been proposed for analyzing mineral reflectance spectra from hyperspectral images. First, unmixing methods [8] are based on endmembers which can be chosen in a database or retrieved from the image. Also, classification algorithms identify the dominant mineral based on similarity measures with reference spectra [9]. These solutions as well as those based on machine learning, rely on reference spectral libraries such as the USGS spectral library [10], which includes more than 1200 reflectance spectra of various minerals. However, these databases are scarce and the number of spectra per mineral is often low. Thus, specific databases have to be created to analyze the physical properties of minerals. Moreover, the spectral signatures may vary depending on the acquisition sensor and the measurement conditions.

Then, knowledge-based approaches, such as spectral indices [11], Tetracoder©[12], Wavelength Mapper [13] and Quantools [14], aim to analyze specific spectral features of the mineral reflectance spectrum. Spectral features yield quantitative information that can be used to identify the mineral. However, they often need prior initialization or expert knowledge and are ill-suited to analyze the full VNIR and SWIR ranges.

Finally, spectral deconvolution approaches rely on a physically-based reflectance model, *e.g.*, the Modified Gaus-

R. Rialland and R. Marion are with CEA/DAM/DIF, F-91680 Bruyères-le-Châtel, France (e-mail: {ronan.rialland,rodolphe.marion}@cea.fr).

C. Soussen is with CentraleSupélec at Laboratoire des Signaux et Systèmes (Université Paris-Saclay-CNRS-CentraleSupélec), F-91190 Gif-sur-Yvette, France (e-mail: charles.soussen@centralesupelec.fr).

V. Carrère is with Université de Nantes at Laboratoire de Planétologie et Géodynamique (UMR 6112), F-44300 Nantes, France (e-mail: veronique.carrere@univ-nantes.fr).

sian Model (MGM) [15] or the Exponential Gaussian Optimization (EGO) model [16], [17]. The deconvolution of mineral reflectance spectra has undergone several improvements over the years with increasingly automatic algorithms [18]–[21] and various applications from iron absorption estimation in the VNIR range [15], [22] to minerals with various chemical compositions such as phyllosilicates [23], [24]. Among them, the Automatized Gaussian Model (AGM) proposed by the authors in [2], is based on the EGO model and is able to deal with both VNIR and SWIR ranges in a same process. The estimated parameters can be used for identification and characterization of a large variety of minerals.

In this paper, we first review the EGO model and the AGM procedure (section II). Then, several substantial improvements, such as a new continuum removal procedure or the use of a greedy algorithm for absorption pre-estimation, are proposed to yield a flexible and automatic procedure (section III). The method is validated on synthetic data (section IV) and then applied to various experimental spectra related to laboratory measurements and hyperspectral images (section V).

II. EGO MODEL

A. Presentation

The EGO model [16], [17] aims at decomposing the logarithm of a mineral reflectance spectrum ρ as a continuum c and a sum of N asymmetric (modified) Gaussian features:

$$\ln \rho\left(\lambda, \boldsymbol{\theta}\right) = c\left(\lambda, \boldsymbol{\theta}_c\right) - \sum_{i=1}^{N} G\left(\lambda, \boldsymbol{\theta}_{G_i}\right) + n(\lambda) \quad (1)$$

$$= \Phi(\lambda, N, \boldsymbol{\theta}_c, \boldsymbol{\theta}_G) + n(\lambda) \quad (2)$$

where λ is the wavelength and $n(\lambda)$ is the noise process. The generic notation $\Phi(\lambda, N, \boldsymbol{\theta}_c, \boldsymbol{\theta}_G)$ expresses the dependency upon both continuum and Gaussian parameters, respectively $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}_G = \{\boldsymbol{\theta}_{G_1}, \ldots, \boldsymbol{\theta}_{G_N}\}$. Bold notations refer to a set of parameters.

The continuum part models the overall shape of the spectrum. The drop-off toward the ultraviolet (below 400 nm), denoted uv, is essentially due to iron whereas the one in the SWIR (beyond 2500 nm), denoted $water$, mostly depends on water and OH absorption. The continuum is then expressed as

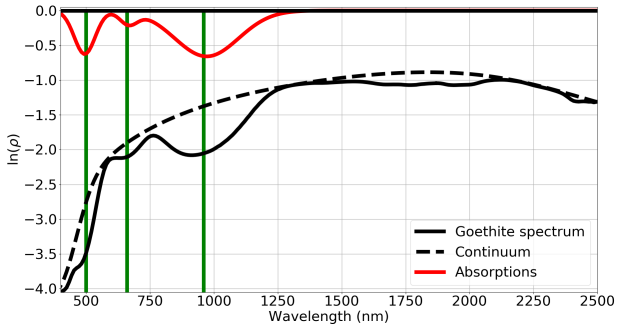
$$\begin{aligned} c\left(\lambda, \boldsymbol{\theta}_c\right) = & -c_0 - c_1\lambda^{-1} - G_{uv}\left(\lambda, s_{uv}, \mu_{uv}, \sigma_{uv}\right) \\ & - G_{water}\left(\lambda, s_{water}, \mu_{water}, \sigma_{water}\right). \end{aligned} \quad (3)$$

G_{uv} and G_{water} are two standard Gaussians with respective amplitudes s_{uv} and s_{water}, positions μ_{uv} and μ_{water} and widths σ_{uv} and σ_{water}. They model the drop-offs on the edges of the spectrum. The offset c_0 and slope c_1 deal with the overall shape. Thus, $\boldsymbol{\theta}_c = \left\{c_0, c_1, s_{uv}, \mu_{uv}, \sigma_{uv}, s_{water}, \mu_{water}, \sigma_{water}\right\}$ gathers the continuum parameters.

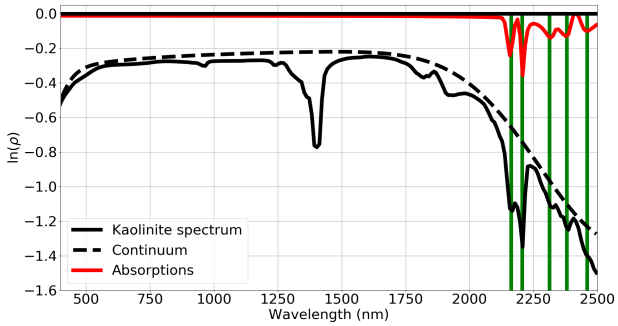
EGO Gaussians model the wide variety of encountered absorption shapes in the VNIR and SWIR ranges. They are defined as asymmetric functions:

$$G\left(\lambda, \boldsymbol{\theta}_{G_i}\right) = s_i \exp\left(-\frac{1}{2}\frac{\left(\lambda - \mu_i\right)^2}{\left(\sigma_i - k_i\left(\lambda - \mu_i\right)\right)^2}\right) \quad (4)$$

with s_i the amplitude, μ_i the position, σ_i the width and k_i the asymmetry of the absorption. Thus, $\boldsymbol{\theta}_{G_i} = \{s_i, \mu_i, \sigma_i, k_i\}$ gathers the parameters of the i-th Gaussian. The EGO model is illustrated in Fig. 1 for goethite and kaolinite spectra with continuum and absorption features estimated by the solution proposed in this paper (only the EGO Gaussians corresponding to the absorptions used to identify the mineral are shown).



(a)



(b)

Fig. 1: Logarithm of goethite (a) and kaolinite (b) reflectance spectra from the USGS spectral library [10]. The estimated continuums are shown with dashed lines and the estimated absorptions in red. The vertical green lines indicate the positions of the theoretical absorptions used to identify the mineral.

B. Limits and identifiability

An attractive feature of the EGO model is that it is independent of site or measurement conditions. Therefore, it has been used in several studies [18], [21], [24]. It provides information on both chemical and physical properties of the surface, and has proven to be one of the most powerful tools for deconvolution of the reflectance spectrum of binary or ternary mixtures as well as overlapping absorption features [2]. However, the EGO model may not be identifiable, since some spectral features may be accurately adjusted by different sets of EGO parameters, as discussed hereafter.

The physics of the continuum remains not fully understood and several continuum models have been proposed, *e.g.*, a convex hull [25], a second-order polynomial [19], a linear spline interpolator [18]. They all lead to errors on the continuum removed spectrum. In the case of the EGO model, we may encounter the following ones. First, the Gaussians G_{uv} and G_{water} may not perfectly adjust the edges of the spectrum which may yield artifacts, similar in shape to EGO

Gaussians. Then, the VNIR part of the spectrum is modeled by the parameter c_1 and G_{uv} which both influence the drop-off at the edge of the spectrum. Note also that the range of variation of the parameter c_1 is large, from 0 to $\sim 10^3$ nm, since c_1 is weighted by λ^{-1} in (1). Finally, EGO Gaussians of low amplitude, wide and located at the center of the spectrum or strongly asymmetric can reduce the estimated value of c_0.

Moreover, a sum of EGO Gaussians may be confused with a single EGO Gaussian. One may note that some authors have introduced a parameter to take the saturation of the absorption bands into account [16]. However, the resulting shape of a saturated EGO Gaussian may also be confused with a sum of two Gaussians. For simplicity reasons, the saturation parameter is not considered here. Furthermore, large values of the asymmetry parameter k may create unrealistic absorption shapes. Bounds based on experiments will be set to avoid this problem. Finally, the number of Gaussians N may lead to overfitting when absorptions are close in position.

To limit these identifiability issues, most semi-automatic methods require initialization based on priors [16]–[19], [24] and bounds [22]. In this paper, we aim at estimating the full set of EGO parameters N, $\boldsymbol{\theta}_c$ and $\boldsymbol{\theta}_G$. The challenges are to adaptively estimate the model order, *i.e.*, the number of unknown parameters, related to the unknown number of Gaussians N, and to automatically initialize the EGO parameters with mineralogically coherent values.

C. Spectral deconvolution issues

The proposed AGM procedure includes continuum removal and absorption estimation steps. Let us review the main technical challenges.

Continuum removal may be carried out using parametric or non-parametric methods. The latter usually rely on convex shape assumption, which makes it possible to apply geometric methods [18]. This approach is fast and easy to apply to various mineral spectra but can be sensitive to noise [26]. Parametric methods are model dependent and suffer from parameter initialization issues [16], [19], [21], [27]. Hereafter, we make use of a parametric method, based on the EGO continuum model, with adapted initialization of parameters $\boldsymbol{\theta}_c$.

Absorption estimation deals with two objectives: determine the number of Gaussians N and estimate the EGO Gaussian parameters $\boldsymbol{\theta}_G$. Several solutions have been proposed such as, *e.g.*, spectral derivatives [20], [21] and the Continuous Wavelet Transform [28], [29]. The former retrieves the local minima of the spectrum but is sensitive to noise and may over-estimate the number of Gaussians. The latter aims to separate the absorptions from the noise. However, the related parameters are difficult to set, because they strongly depend on the mineral reflectance spectrum, and the number of Gaussians is often under-estimated. Note also that these methods do not estimate the asymmetry parameter. On the contrary, our method retrieves the EGO (asymmetric) Gaussian features. Importantly, the proposed approach includes an adaptive estimation of the number of EGO Gaussians depending on the noise statistics.

III. Improved AGM procedure

A. Overview

The improved AGM procedure takes as input a mineral reflectance spectrum $\rho(\lambda)$ and the associated noise covariance matrix $\boldsymbol{\Sigma}$. It is composed of three steps: *(i)* continuum pre-estimation, *(ii)* absorption pre-estimation, and *(iii)* joint optimization. Continuum pre-estimation is carried out using a non-linear least-squares formulation, with initial values of $\boldsymbol{\theta}_c$ defined beforehand. The resulting pre-estimated continuum can then be subtracted from the data spectrum:

$$a(\lambda) = c(\lambda, \boldsymbol{\theta}_c^{pre}) - \ln \rho(\lambda) \tag{5}$$

where $\boldsymbol{\theta}_c^{pre}$ and $a(\lambda)$ respectively refer to the pre-estimated continuum parameters and the absorption part, that is, the sum of EGO Gaussians in (1). Absorption pre-estimation then automatically retrieves the parameters of the EGO Gaussians, noted $\boldsymbol{\theta}_G^{pre}$ from $a(\lambda)$. These continuum and absorption estimates are refined during the joint optimization step, which boils down to a least-squares post-processing stage. The improved AGM procedure is summarized in Fig. 2.

In the following, the noise is assumed to be zero-mean, Gaussian and independent from one spectral band to another. However, the noise variance σ_λ^2 is known to depend on the spectral band λ. The noise covariance matrix $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_\lambda^2)$ is thus diagonal and can be either calculated knowing the sensor characteristics or estimated from the data [30].

The three steps of the AGM procedure are detailed in the next subsections. The column vector gathering the reflectance values $\rho(\lambda)$ for every λ is denoted by $\boldsymbol{\rho} \in \mathbb{R}^{N_\lambda}$, with N_λ the number of wavelengths. Similarly, the vectors $\mathbf{c}(\boldsymbol{\theta}_c)$, $\mathbf{a}(\boldsymbol{\theta}_G)$ and $\boldsymbol{\sigma}$ gather the values of $c(\lambda, \theta_c)$, $a(\lambda, \boldsymbol{\theta}_G)$ and σ_λ, respectively, for all λ.

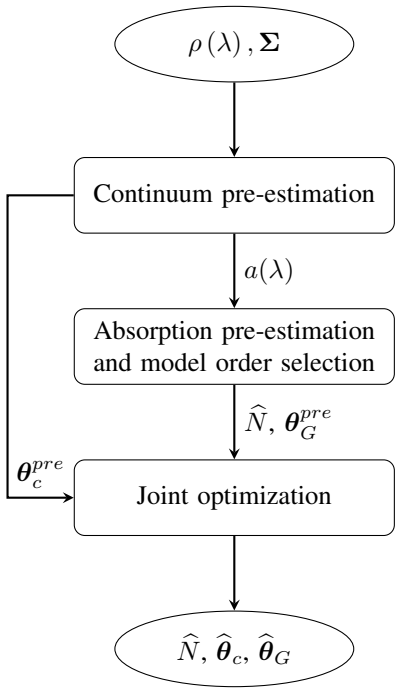


Fig. 2: Flowchart of the improved AGM procedure.

B. Continuum pre-estimation

The following least-squares formulation corresponds to the maximum likelihood estimation of parameters $\boldsymbol{\theta}_c$:

$$\min_{\boldsymbol{\theta}_c} \|\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\mathbf{c}-\ln\boldsymbol{\rho}\left(\boldsymbol{\theta}_c\right)\right)\|^2 \;\text{ s.t. }\; \begin{cases}\mathbf{c}\left(\boldsymbol{\theta}_c\right)-\ln\boldsymbol{\rho}\geq\alpha\boldsymbol{\sigma}\\ \boldsymbol{\theta}_c^{min}\leq\boldsymbol{\theta}_c\leq\boldsymbol{\theta}_c^{max}\end{cases} \tag{6}$$

where $\|\,.\,\|$ refers to the Euclidean norm, and the compact writing $\mathbf{c}\left(\boldsymbol{\theta}_c\right)-\ln\boldsymbol{\rho}\geq\alpha\boldsymbol{\sigma}$ refers to N_λ constraints, one per wavelength. In a nutshell, the latter constraints aim to impose that the absorption $a(\lambda)$ defined in (5) is non-negative valued, that is, the amplitude of the EGO Gaussians are positive. The coefficient α is set empirically to adapt the tolerance to the noise standard deviation. Bounds $\boldsymbol{\theta}_c^{min}$ and $\boldsymbol{\theta}_c^{max}$ are defined in subsection III-E. Other constraints may be used to impose a specific continuum shape, *e.g.*, concavity constraints. However, such constraint is not always suited to deal with a large variety of minerals. The least-squares problem (6) is solved using the Constrained Optimization BY Linear Approximation (COBYLA) algorithm [31].

Since (6) is a non-linear least-squares problem, the numerical solution $\boldsymbol{\theta}_c^{pre}$ is likely to be a local minimizer, which depends on the initial value of $\boldsymbol{\theta}_c$, denoted $\boldsymbol{\theta}_c^{start}$. The initialization strategy is illustrated in Fig. 3 for goethite. The offset c_0 is set as the maximum value of the spectrum and the continuum slope c_1 is set to 0 [21]. Starting values of the Gaussians G_{uv} and G_{water} are computed similarly in the VNIR and SWIR spectral ranges. Their positions are set to predefined values $\mu_{uv}=200$ nm and $\mu_{water}=2800$ nm [27]. Then, the lines l_1 and l_2 are computed between respectively the values of the spectrum at the first and last wavelength and the maximum values of the spectrum in the VNIR, at $\lambda_{\rho_{VNIR}^{max}}$, and in the SWIR, at $\lambda_{\rho_{SWIR}^{max}}$. The amplitudes s_{uv} and s_{water} are set to the values of l_1 and l_2 at $\mu_{uv}=200$ nm and $\mu_{water}=2800$ nm with c_0 removed. Finally, the widths σ_{uv} and σ_{water} are computed according to the settings of Fig. 3. This strategy is adapted to the various continuum shapes of the minerals of interest, as G_{uv} and G_{water} only influence their own spectral range, *i.e.*, the VNIR for G_{uv} and the SWIR for G_{water}.

C. Pre-estimation of absorption characteristics

Once the continuum is pre-estimated, (5) yields the absorption signal $a(\lambda)$, which rereads

$$a(\lambda)=\sum_{i=1}^{N}G\left(\lambda,\boldsymbol{\theta}_{G_i}\right)-n(\lambda)+e(\lambda) \tag{7}$$

according to (1), where $e(\lambda)$ refers to the estimation error in the continuum pre-estimation step. Absorption estimation consists in decomposing $a(\lambda)$ into a weighted sum of Gaussians with positive weights. This problem can be effectively solved using greedy algorithms, which select the EGO Gaussians one by one in a large dictionary containing all possible EGO Gaussians. This greedy approach is detailed hereafter. Moreover, an information theoretic criterion will be used to adaptively estimate the number of selected Gaussians.

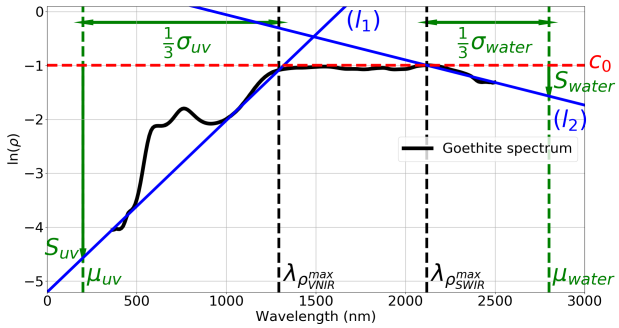


Fig. 3: Computation of the initial continuum parameters $\boldsymbol{\theta}_c^{start}$ from a goethite reflectance spectrum, in black, from the USGS spectral library [10], sample: `splib07b_Goethite_GDS134_ASDFRb_AREF`. The lines l_1 and l_2 are drawn in blue, and the maximum of the spectrum c_0 in dashed red. The wavelengths $\lambda_{\rho_{VNIR}^{max}}$ and $\lambda_{\rho_{SWIR}^{max}}$ of the maximum of the spectrum in the VNIR and SWIR are in dashed black. The parameters of G_{uv} and G_{water} are in green.

1) Estimation of EGO Gaussian parameters: Starting from (7), we adopt an optimization approach akin to (6), which consists of minimizing the negative log-likelihood

$$\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\mathbf{a}-\sum_{i=1}^{N}G\left(\lambda,\boldsymbol{\theta}_{G_i}\right)\right)\right\|^2 \tag{8}$$

with respect to the EGO Gaussian parameters $\boldsymbol{\theta}_G$. The difficulty of this optimization problem is twofold. First, this is a non-linear least-squares problem, thus the cost function is likely to be multimodal, *i.e.*, with many local minimizers. Moreover, the number of Gaussians N is unknown.

We propose to solve the latter optimization problem for consecutive values of N, yielding absorption decompositions for many N. Then, a model order selection rule [32] is applied to select one particular decomposition among all. An attractive feature of greedy algorithms is that they provide solutions to the minimization problems (8) for $N=1,\ldots,N_{max}$ with a single run. N and N_{max} are the current iteration of the algorithm and the total number of iterations, respectively. The absorption decomposition found for a given N is used as initialization for the search of the decomposition with $N+1$ Gaussians at the next iteration.

The greedy approach is based on the creation of a very large dictionary $\mathbf{G}\in\mathbb{R}^{N_\lambda\times N_{atom}}$ containing all possible spectral features. Each column (also called atom) of $\mathbf{G}$ refers to a specific EGO Gaussian with predefined shape parameters (μ,σ,k), and magnitude $s=1$. The EGO Gaussian model (4) thus reads as $s_i\,\mathbf{g}_i$ where $\mathbf{g}_i$ refers to a column of $\mathbf{G}$ related to the i-th absorption. In turn, the sum of EGO Gaussians in (8) reads as a matrix-vector product $\mathbf{Gs}$ where $\mathbf{s}\in\mathbb{R}^{N_{atom}}$ are the (non-negative) magnitudes. Thus, the minimization of the residual error (8) rereads:

$$\min_{\mathbf{s}}\|\boldsymbol{\Sigma}^{-1/2}(\mathbf{a}-\mathbf{Gs})\|^2 \;\text{ s.t. }\; \begin{cases}\mathbf{s}\geq\mathbf{0}\\ \|\mathbf{s}\|_0\leq N\end{cases} \tag{9}$$

Algorithm 1 Solve problem (9) for $N = 1, \dots, N_{max}$. The support S gathers the indices of the nonzero elements in $\mathbf{s}$. $\mathbf{g}_i$ refers to the i-th column of $\mathbf{G}$. $\mathbf{G}_S$ refers to the submatrix gathering the columns of $\mathbf{G}$ indexed by S.

inputs: $\mathbf{a}$, $\mathbf{\Sigma}$, $\mathbf{G}$, N_{max}
outputs: $\mathbf{s}^N$, r^N **for all** N
  $\mathbf{r} \leftarrow \mathbf{\Sigma}^{-1/2}\mathbf{a}$
  $\mathbf{s} \leftarrow \mathbf{0}$
  $S \leftarrow \emptyset$
  for $N = 1$ **to** N_{max} **do**
    $\ell \leftarrow \arg\max_i \{\mathbf{g}_i^T \mathbf{\Sigma}^{-1/2}\mathbf{r} \,/\, \|\mathbf{\Sigma}^{-1/2}\mathbf{g}_i\|, i \notin S\}$
    $S \leftarrow S \cup \{\ell\}$
    $\mathbf{s}(S) \leftarrow \arg\min_{\mathbf{z}} \{\|\mathbf{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{G}_S\mathbf{z})\|^2,\ \mathbf{z} \geqslant \mathbf{0}\}$
    $\mathbf{s}^N \leftarrow \mathbf{s}$
    $\mathbf{r}^N \leftarrow \mathbf{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{G}\mathbf{s}^N)$
  end for

where $\|\mathbf{s}\|_0$ counts the number of non-zero entries in vector $\mathbf{s}$ and N is the number of absorption features embedded in a mineral reflectance spectrum.

2) NNOMP, a greedy algorithm: Non-Negative Orthogonal Matching Pursuit (NNOMP) [33], [34] is a greedy search algorithm dedicated to solving problem (9) for consecutive $N = 1, \dots, N_{max}$. As shown in Algorithm 1, NNOMP gradually selects atoms in $\mathbf{G}$ in order to refine the approximation $\mathbf{a} \approx \mathbf{G}\mathbf{s}$. Initially, no atom is selected, then $\mathbf{s} = \mathbf{0}$. At each iteration, a new atom of $\mathbf{G}$ is added to the current support S gathering the previously selected atoms. Their amplitudes are then adjusted using a Non-Negative Least-Squares solver [34]. The residual $\mathbf{r} = \mathbf{\Sigma}^{-1/2}(\mathbf{a} - \mathbf{G}\mathbf{s})$ is updated accordingly.

The atom selection rule of NNOMP aims at decreasing at most the norm of the residual $\mathbf{r}$. Thus, in practice, the main absorptions of the spectrum are retrieved in the first iterations, then smaller absorptions are gradually retrieved.

3) Model order selection: Model order selection aims at automatically selecting a solution $\mathbf{s}^N$ among the outputs $\mathbf{s}^N$, $N = 1, \dots, N_{max}$. Many rules such as the Akaike information criterion tend to overestimate the true number of features, because they were designed based on an asymptotic information theory analysis, where the number of observations tend to infinity [32]. Minimum Description Length (MDL) criteria are easily adaptable to greedy algorithms [35]. Moreover, they were adapted to short data records, where the number of parameters is moderately smaller than the data size as in our case [36]. In the case of known noise statistics, they read:

$$\widehat{N} = \arg\min_N \left(\ln \|\mathbf{r}^N\| + \frac{\ln N_\lambda \, (N+1)}{N_\lambda - N - 2} \right). \qquad (10)$$

This criterion is thus evaluated at each iteration of NNOMP. Then, the NNOMP output $\mathbf{s} = \mathbf{s}^N$ for $N = \hat{N}$ is selected.

4) Dictionary creation and $\boldsymbol{\theta}_G$ pre-estimation: The dictionary $\mathbf{G}$ gathers a large number of EGO Gaussians computed for a set of position (μ), width (σ) and asymmetry (k) parameters. To do so, each EGO parameter is discretized over a pre-defined grid whose resolution is related to the size of the dictionary. Fine grids induce a greater precision at the cost of increased memory storage and computational time.

The NNOMP algorithm yields the number $\hat{N}$ of selected atoms, their column indices $\mathbf{g}_i$ in the dictionary $\mathbf{G}$ and their magnitudes s_i. Since there is a one-to-one correspondence between a column of $\mathbf{G}$ and a set of parameters (μ, σ, k), the parameters related to the i-th selected atom, *i.e.*, $\boldsymbol{\theta}_{G_i}^{pre} = \{s_i, \mu_i, \sigma_i, k_i\}$ can be retrieved from the knowledge of the location of $\mathbf{g}_i$ in the dictionary $\mathbf{G}$ and from the magnitude s_i.

D. Joint optimization

Once the continuum parameters $\boldsymbol{\theta}_c^{pre}$, the number of Gaussians $\widehat{N}$ and the associated parameters $\boldsymbol{\theta}_G^{pre}$ have been pre-estimated, a joint optimization process is performed to refine the continuum and absorption estimates, whose accuracy depends on the discretization of the parameters in the dictionary $\mathbf{G}$. We define the following non-linear least-squares minimization problem with $N = \widehat{N}$:

$$\min_{\boldsymbol{\theta}_c, \boldsymbol{\theta}_G} \left\| \mathbf{\Sigma}^{-\frac{1}{2}} \left(\ln \boldsymbol{\rho} - \mathbf{c}\left(\boldsymbol{\theta}_c\right) + \sum_{i=1}^{N} \mathbf{G}\left(\boldsymbol{\theta}_{G_i}\right) \right) \right\|^2. \qquad (11)$$

The problem is solved by the Trust Region Reflective algorithm [37] using initial parameters $\boldsymbol{\theta}_c^{pre}$ and $\boldsymbol{\theta}_G^{pre}$.

E. Algorithm parametrization

This section summarizes the parameter settings within the proposed AGM procedure. We first state the constraints related to the continuum pre-estimation step and then the parameters and constraints used in the absorption pre-estimation step.

Bounds on the continuum parameters $\boldsymbol{\theta}_c$ are defined in (6) to ensure that the range of EGO parameters is consistent with their physical interpretation. As $\ln \rho(\lambda)$ is negative valued, the parameters c_0, c_1, s_{uv} and s_{water} in (3) are imposed to be positive. Also, this ensures the drop-offs of the continuum to appear on the edges of the spectrum. The Gaussians G_{uv} and G_{water} are supposed to be centered outside the spectral range of the acquisition sensors with $\mu_{uv} \in \left[0 - \lambda^{min}\right]$ nm and $\mu_{water} \in [\lambda^{max} - 3000]$ nm. To avoid unrealistic values, lower (0 nm) and upper (3000 nm) bounds are set for both μ_{uv} and μ_{water}. Widths of G_{uv} and G_{water} are unconstrained. The tolerance α in (6) deals with the noise standard deviation. α is set to 3 in noisy cases, and to 0 in the noise-free cases.

In Algorithm 1, the maximum number of selected atoms N_{max} is set to 20 to limit the possible number of absorptions. Since our estimation of N is adaptive, $\hat{N} = N_{max}$ is rarely reached in practice. Indeed, N often ranges between 1 and 5 for mineral reflectance spectra. To create the dictionary $\mathbf{G}$, we define two grids adapted to the various absorption shapes in, respectively, the VNIR and SWIR ranges. VNIR absorptions tend to be broad whereas SWIR absorptions are narrow and occasionally asymmetric. In the VNIR, k_i is set to 0, as VNIR absorptions are generally symmetric. The width σ_i varies between 30 and 380 nm with a step $\frac{p}{2}$, where p is the wavelength sampling step of the sensor (for AVIRIS, $p \sim 10$ nm). The latter minimum width is set experimentally to avoid selecting atoms of small amplitudes that do not identify with

an absorption. The maximum width corresponds to maximum ferrous goethite absorption at 920 nm, measured on goethite spectra from the USGS spectral library [10]. The positions vary between λ^{min} and 1300 nm, with step $\frac{p}{2}$. In the SWIR, k_i varies between -0.2 and 0.2 to avoid the creation of unrealistic EGO Gaussians with a step of 0.05. These values were chosen after several experiments on mineral spectra with asymmetric absorptions in the SWIR. The width varies between 5 and 45 nm with a step equal to $\frac{p}{2}$. The maximum width is fixed experimentally, based on the application of the improved AGM procedure on an alunite spectrum. The positions vary between 1300 nm and λ^{max} with a step of $\frac{p}{10}$.

IV. Validation on synthetic data

In this section, the improved AGM procedure is validated on noise-free and noisy synthetic data.

A. Synthetic data description and error metrics

Three synthetic spectra are generated using the EGO model in order to represent various continuum and absorption shapes (amplitude, width, asymmetry). These spectra are shown in Fig. 4. Their parameters are given in Tab. I. Spectra 1 and 2 have three absorptions and spectrum 3 four. The spectra are convolved to the AVIRIS spectral response function. The objective is to focus on difficult cases with broad and shallow absorptions in the VNIR (like for goethite and hematite), close and narrow absorptions in the SWIR (like for kaolinite), asymmetrical absorptions (like for alunite and calcite), and to assess the impact of the continuum shape and the Gaussians used to model the typical drop-offs in the UV and SWIR.

Hereafter, the discrepancy between a synthetic (noise-free) spectrum $\boldsymbol{y}^{syn}$ and its recovered version $\boldsymbol{y}^{est}$ is assessed using the goodness of fit, expressed in dB:

$$r = 10\ \log_{10} \frac{\|\boldsymbol{y}^{syn}\|^2}{\|\boldsymbol{y}^{syn} - \boldsymbol{y}^{est}\|^2}. \tag{12}$$

This indicator is computed for each step of the improved AGM procedure, yielding ratios r_1, r_2 and r_3 between the synthetic and estimated continuums $\boldsymbol{c}^{syn}$ and $\boldsymbol{c}^{est}$, absorption spectra $\boldsymbol{a}^{syn}$ and $\boldsymbol{a}^{est}$, and spectra $\ln \boldsymbol{\rho}^{syn}$ and $\ln \boldsymbol{\rho}^{est}$, respectively. Furthermore, we introduce metrics to compare the synthetic parameters $\boldsymbol{\theta}^{syn}$ and their estimates $\boldsymbol{\theta}^{est}$. For the absorption positions μ, we use the signed error $\delta_\mu = \mu_{est} - \mu_{syn}$. As for other EGO parameters θ_i, the error is expressed as the normalized ratio

$$\Delta_{\theta_i} = 100 \left| \frac{\theta_i^{est} - \theta_i^{syn}}{\theta_i^{syn}} \right|. \tag{13}$$

B. Algorithm validation in noise-free cases

Each synthetic spectrum is deconvolved with the improved AGM procedure, see Fig. 4. For noise-free cases, the noise covariance $\boldsymbol{\Sigma}$ is set to the identity matrix.

In the continuum pre-estimation step (Fig. 4(d-f)), the goodness of fit between the pre-estimated continuum and the synthetic one is $r_1 \sim 30$ dB for the three spectra. The errors on the continuum amplitudes are lower than 10%. They mostly occur in the spectral ranges of broad absorptions (*e.g.*, at 960 nm for spectrum 1), at central wavelengths (*e.g.*, between 1300 and 2000 nm for spectrum 1), and near the edges of the spectrum (*e.g.*, for wavelengths larger than 2450 nm for spectrum 1). These errors are related to the presence of broad absorptions in the spectrum and to the fact that G_{uv} and G_{water} are centered outside the $[400 - 2500]$ nm range, where data are missing. These errors may create three kinds of artifacts in the absorption spectra. First, the amplitudes and shapes of broad absorptions can be impacted (*e.g.*, absorption at 960 nm of spectrum 1 and at 2160 nm of spectrum 2). Notably, errors in the VNIR are larger as the continuum is modeled by both c_1 and G_{uv}. Then, small and broad false absorptions may appear at central wavelengths. Finally, small and narrow false absorptions may appear on the edges of the spectra. Note that such artifacts are not specific to our continuum removal procedure. For example, the Clark and Roush procedure [25], used in the GMEX (Guide for Mineral EXploration) [38] also creates artefacts at both edges of the spectral range.

The absorption pre-estimation step, applied to the *synthetic* absorption spectra (*i.e.*, simulated without continuum), is illustrated in Fig. 5(a-c). First, r_2 is close to 20 dB for the three spectra as broad and overlapping absorptions may be slightly shifted (*e.g.*, $\sim$ 20 nm for the absorption at 960 nm of spectrum 1, $\sim$ 10 nm for the absorption at 2160 nm of spectrum 2) from their synthetic positions. When overlapping absorptions are similar in amplitude and shape, as for the absorption doublet at 2162 and 2206 nm of spectrum 3, three EGO Gaussians are recovered including a false absorption. On the contrary, isolated absorptions are well retrieved with errors lower than 0.5% for all parameters.

The absorption pre-estimation step, applied to the absorption spectra after the continuum pre-estimation step (*i.e.*, with continuum artifacts) is illustrated in Fig. 4(g-i). Compared to the results of Fig. 5, the number of false EGO Gaussians is increased due to the artifacts, but their amplitudes are small. For example, in spectrum 1, two false absorptions appear around 1900 and 2500 nm and are fitted by four EGO Gaussians. Note that, these false EGO Gaussian may be easily handled in a mineral identification procedure as they do not correspond to known mineral absorptions.

Finally, the goodness of fit r_3 is around 60 dB for the three spectra, which confirms that the recovered spectrum $\boldsymbol{\rho}^{est}$ perfectly identifies with $\boldsymbol{\rho}^{syn}$, see Fig. 4(j-l). The estimated continuums are consistent with the shape of the three spectra. Absorptions are retrieved with small errors, especially on their positions. As an example, the maximum position error is reached for the broad and overlapping absorption at 960 nm of spectrum 1 ($\delta_{mu} = 40$ nm). For all other absorptions of the three spectra, $\delta_{mu} < 3$ nm, which is lower than the current spectral resolution of airborne sensors such as AVIRIS ($\sim$ 10 nm). Also, absorption shape parameters are accurately retrieved such as, *e.g.*, the asymmetry of the broad and overlapping absorption at 2165 nm of spectrum 2 ($k = -0.2$). For isolated absorptions (*e.g.*, at 2283 nm for spectrum 1 and at 1760 nm for spectrum 2), errors on the amplitudes, widths and asymmetry are lower than 1%.

	Continuum parameters				Absorption parameters			
	c_0	c_1	G_{uv}	G_{water}	Abs 1	Abs 2	Abs 3	Abs 4
spectrum 1	0.5	500	$\mu = 200$ $\sigma = 250$ $s = 1.2$	$\mu = 2800$ $\sigma = 200$ $s = 1.0$	$\mu = 660$ $\sigma = 40$ $s = 0.1$ $k = 0.0$	$\mu = 960$ $\sigma = 125$ $s = 0.25$ $k = 0.0$	$\mu = 2283$ $\sigma = 7$ $s = 0.4$ $k = 0.2$	
spectrum 2	0.5	0.01	$\mu = 200$ $\sigma = 250$ $s = 1.2$	$\mu = 2800$ $\sigma = 400$ $s = 0.8$	$\mu = 1760$ $\sigma = 12$ $s = 0.3$ $k = 0.0$	$\mu = 2165$ $\sigma = 45$ $s = 0.4$ $k = -0.25$	$\mu = 2324$ $\sigma = 10$ $s = 0.25$ $k = 0.0$	
spectrum 3	0.2	0.01	$\mu = 200$ $\sigma = 250$ $s = 1.2$	$\mu = 2800$ $\sigma = 400$ $s = 1.0$	$\mu = 2162$ $\sigma = 15$ $s = 0.35$ $k = 0.0$	$\mu = 2206$ $\sigma = 17$ $s = 0.45$ $k = 0.0$	$\mu = 2312$ $\sigma = 10$ $s = 0.05$ $k = 0.0$	$\mu = 2380$ $\sigma = 10$ $s = 0.05$ $k = 0.0$

TABLE I: Synthetic EGO parameters.

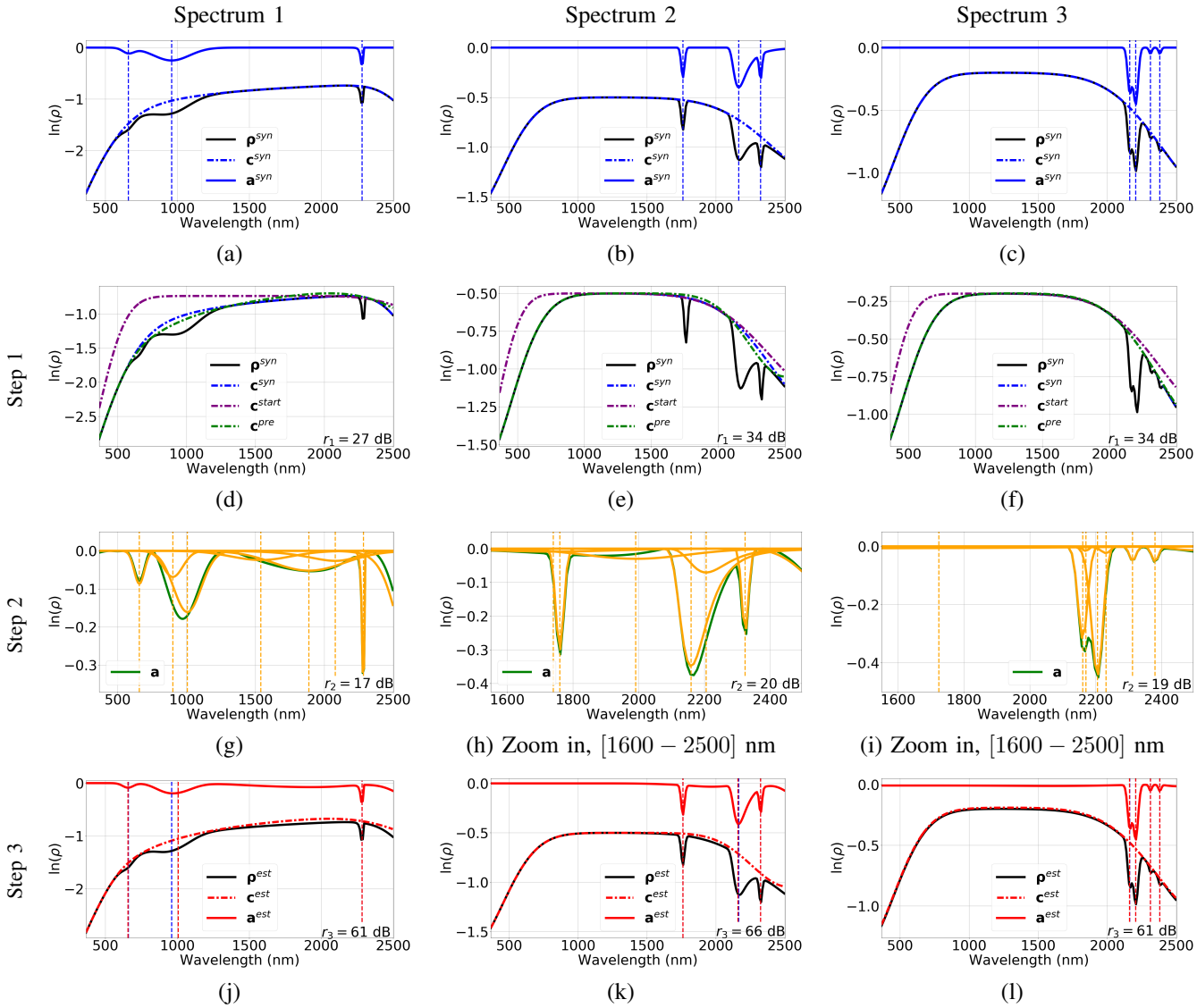
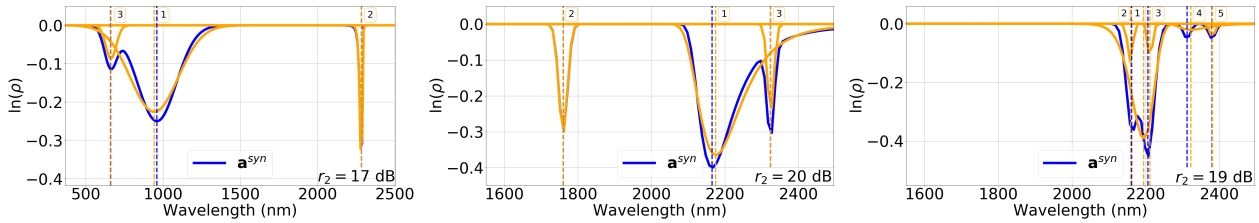
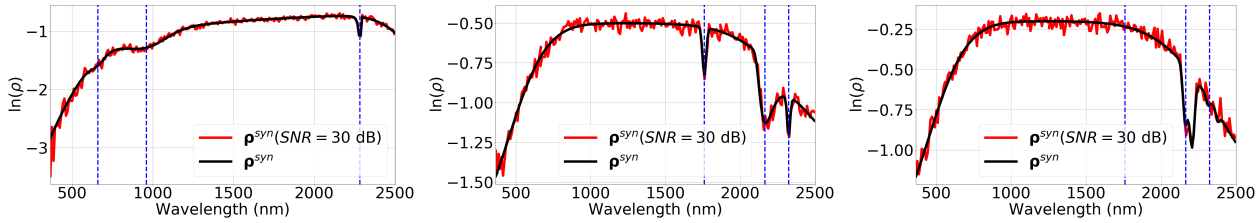


Fig. 4: Results of the improved AGM procedure on three synthetic spectra. (a,b,c) Synthetic spectra $\boldsymbol{\rho}^{syn}$ (solid black curve), absorption $\boldsymbol{a}^{syn}$ (solid blue) and continuum $\boldsymbol{c}^{syn}$ (dashdot blue). The absorption positions are shown with vertical dashed lines. (d,e,f) Starting $\boldsymbol{c}^{start}$ (purple - dashdot) and pre-estimated continuum $\boldsymbol{c}^{pre}$ (green - dashdot). (g,h,i) Absorption spectra after the continuum pre-estimation step (green) and Gaussians selected by NNOMP (orange). Their positions are shown with vertical dashed lines. (j,k,l) Estimated absorption spectra $\boldsymbol{a}^{est}$ and continuum $\boldsymbol{c}^{est}$ (red). For each step, the goodness of fit (r_1, r_2, r_3) is given in dB.



(a) Spectrum 1 (b) Spectrum 2 (c) Spectrum 3

Fig. 5: Synthetic absorption spectra (blue) and Gaussians selected by NNOMP (orange). The order of selection is also indicated. The true absorption positions and their estimates are shown as vertical dashed lines (respectively blue and orange).



(a) Spectrum 1 (b) Spectrum 2 (c) Spectrum 3

Fig. 6: Example of synthetic spectra for SNR $= 30$ dB. Estimated EGO Gaussian positions are shown as vertical dashed lines.

C. Algorithm validation on noisy cases

The three synthetic spectra are corrupted with an additive Gaussian noise of zero-mean, whose variance is identical on each spectral band. The noise variance is expressed as $\sigma^2 = 10^{-\mathrm{SNR}/10} \ln \overline{\rho}$, where $\overline{\rho}$ is the mean of the synthetic spectrum and SNR is the Signal-to-Noise Ratio, expressed in dB. Noise spectra are generated for SNRs ranging from 15 to 70 dB, where the latter SNR corresponds to the noise-free simulation. Fig. 6(a-c) show the three spectra with SNR $= 30$ dB. 100 noise spectra are generated per SNR. For each EGO parameter, we compute the standard deviation, noted ϵ_θ, from the 100 estimates of θ.

Qualitative results are close to those of Fig. 4, therefore they are omitted. As expected, when the SNR increases, the reconstruction accuracy increases. The goodness of fit r_3 ranges from ~ 10 to ~ 60 dB for SNRs equal to 15 and 70 dB, respectively. For SNRs lower than 25 dB, only the broadest and deepest absorptions are retrieved. Moreover, the number of estimated EGO Gaussians is underestimated. Indeed, EGO features of small amplitudes are hidden in the noise, thus they cannot be recovered. For SNR greater than 30 dB, the broadest and deepest absorptions are always retrieved with small errors on their positions (*e.g.*, $\epsilon_\mu = 9$ nm for the absorption at 2160 nm of spectrum 2). Also, the isolated absorptions are fitted by unique EGO Gaussians and the continuum parameters are retrieved with small errors on G_{uv} and G_{water} (*e.g.*, $\epsilon_{\mu_{water}} = 10$ nm and $\epsilon_{\mu_{water}} = 17$ nm for spectrum 2). For SNR greater than 50 dB, all the absorptions, even those with small amplitudes, are fitted and their widths correspond to the expected values (*e.g.*, width error is equal to 0.26 nm for the absorption at 2324 nm of spectrum 2). At last, the asymmetry parameter is retrieved for noise-free data only. It can hardly be identified with precision since a wide range of values yield similar spectral shapes.

V. Application to Real Data

A. AGM procedure for real data

In hyperspectral imaging, spectral ranges around 1400 and 1900 nm may be affected by the main saturated atmospheric water vapor absorptions. Spectral masks (represented as grey areas in the following figures) are thus applied to remove these spectral bands. In order to cope with these missing data, the dictionary creation is slightly adapted within the AGM absorption pre-estimation step, by including dictionary atoms centered in the spectral masks. This allows us to retrieve absorptions inside the spectral masks from available data on the edges of the spectral masks, which avoids the creation of artifacts. Interestingly, the AGM accuracy is not impacted by these spectral masks. In subsections V-B and V-C, spectral masks are applied to laboratory and image spectra.

B. Results on laboratory data

Here, the goal is to analyze the advantages and limits of the method for minerals often encountered in hyperspectral imaging, that entail spectral features difficult to deconvolve. The ability to identify minerals from their estimated EGO parameters is also discussed.

1) Spectral library description and data preparation: We use the USGS spectral library [10], which gathers laboratory VNIR and SWIR reflectance spectra of various materials. Six reference minerals are considered, for which the electronic and vibrational processes are well-known (Tab. II). The continuum and the absorption bands differ, in shape and position (for the absorption), for each mineral spectrum as they belong to different mineralogical classes. In order to obtain results comparable to those of synthetic data, the spectra are convolved to AVIRIS spectral responses. The estimates yielded by the AGM procedure will be numerically compared to the theoretical values in Tab. II, seen as ground truth.

Mineral - Group	Sample	Diagnostic absorption (nm)	Secondary absorption (nm)
Calcite - Carbonate	splib07b_Calcite_GDS304_75-150um_ASDFRb_AREF	2342.5	2156
Dolomite - Carbonate	splib07b_Dolomite_HS102.4B_ASDNGb_AREF	2324	2140
Goethite - Fe-hydroxyde	splib07b_Goethite_GDS134_ASDFRb_AREF	660 960	500
Gypsum - Sulphate	splib07b_Gypsum_HS333.4B_(Selenite)_ASDFRa_AREF	1750	1538 2215
Kaolinite - Phyllosilicate	splib07b_Kaolinite_CM9_ASDNGb_AREF	2162 2206	2312 2380
Nontronite - Smectite	splib07b_Nontronite_NG-1.a_ASDNGb_AREF	2283 660 960	2378

TABLE II: Some minerals of interest in the USGS spectral library [10].

Goethite (α-FeO(OH)) is a Fe-hydroxyde. Its reflectance spectrum exhibits two diagnostic absorptions at 660 and 960 nm, which are attributed to Fe^{3+} electronic processes, transition in the iron cations (Fe^{3+}) from ground state ${}^6A_{1\nu}$ to ${}^4T_{1\nu}$ [39]. The fall-off short of 550 nm is attributed to a conduction band, typical of the trans-opaque iron oxides. We are here evaluating the efficiency of the method to model typically broad absorptions related to iron in the VNIR.

Kaolinite ($Al_2Si_2O_5(OH)_4$) is a phyllosilicate. Its reflectance spectrum exhibits a doublet of diagnostic absorptions at 2162 and 2206 nm. This doublet is related to the Al-OH combination band and is indicative of the two layer dioctahedral structure of kaolinite [40]–[42]. We are here evaluating the ability of the method to estimate a doublet of narrow absorptions, with small amplitudes, in the SWIR.

Nontronite ($Na_{0.33}Fe_2^{3+}(Si,Al)_4O_{10} \bullet nH_2O$) is a smectite. Its reflectance spectrum exhibits three diagnostic absorptions, at 660, 960 and 2283 nm. As for goethite, the 660 and 960 nm absorptions are attributed to electronic transition in the Fe^{3+}ion. The 2283 nm absorption is a Fe-OH combination band [42], [43]. The objective here is to test the accuracy of the method in both VNIR and SWIR ranges.

Gypsum ($CaSO_4 \bullet 2H_2O$) is a sulphate. Its reflectance spectrum exhibits one diagnostic absorption, at 1750 nm, whose origin is controversial. It is either associated to H_2O combination overtones [44], to $OH-$ or H_2O bending, stretching and rotational overtones and/or $S-O$ bending overtones [45]. It has been partly associated to H_2O because various experiments involving dehydration of gypsum demonstrated an attenuation of this specific absorption with increasing temperature and the fact that it became narrower and less structured starting at 100°C [44]. The objective here is to evaluate the capacity of the method to detect a diagnostic absorption when it is close to the "atmospheric" spectral mask.

Calcite ($CaCO_3$) and dolomite ($CaMg(CO_3)_2$) are carbonates. Their reflectance spectra exhibit a diagnostic absorption, respectively at 2342 and 2324 nm, and a secondary one, respectively at 2156 and 2140 nm, which can be interpreted as overtones and combinations of bending and stretching vibrations of the CO_3^{2-} ion [46]–[48]. The slight shift for dolomite is due to the presence of Mg in addition to Ca associated with the CO_3^{2-} ion. The objective here is to evaluate the capacity of the method to discriminate calcite from dolomite despite their close diagnostic absorption features. The influence of the asymmetry factor on the estimate of absorption characteristics is also evaluated.

Note that a post-processing step on the AGM results could be applied to remove unrealistic or overfitted EGO Gaussians (represented in blue in the following figures).

2) VNIR absorptions: goethite ferric iron absorptions: Fig. 7 synthesizes the results for goethite. We focus here on the continuum shape, the diagnostic absorptions at 660 and 960 nm and the secondary absorption at 500 nm.

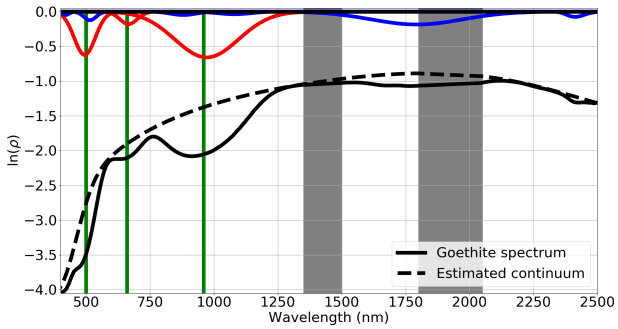


Fig. 7: Goethite spectrum (black), estimated continuum (black-dashed), diagnostic and secondary Gaussians (red), other Gaussians (blue). The true positions of the diagnostic and secondary estimated Gaussians are shown as vertical lines.

First, the continuum is correctly estimated. It fits the overall shape of the spectrum in the SWIR, where no absorption of mineralogical interest is present. In the VNIR, the fit is more complex as both continuum and absorptions have to be taken into account. We thus note that G_{uv} is close to the edge of the spectrum with a small width ($\mu_{uv} = 402$ nm, $\sigma_{uv} = 65$ nm) and that c_1 is large ($c_1 = 1110$). In this case, the fit with EGO Gaussians of the absorption at 500 nm mainly depends on G_{uv} while the ones at 660 and 960 nm depend on c_1. However, the estimated continuum, in the VNIR, lays above the spectrum and its shape is consistent with the results on synthetic data, see spectrum 1 in section IV-B.

As expected, the absorption at 960 nm is fitted by a broad EGO Gaussian whose position is slightly shifted ($\mu = 971$ nm). Since the EGO Gaussian is broad, this shift impacts the absorption at 660 nm and tends to adjust both absorptions at once. The absorption at 660 nm is thus fitted by an EGO Gaussian of reduced amplitude. In addition, three Gaussians (at 880, 997 and 1079 nm) of low amplitudes (respectively 0.05, 0.03 and 0.04) are selected by NNOMP to reduce the

reconstruction error in this area. Finally, the absorption at 500 nm is difficult to adjust because its shape strongly depends on G_{uv}. Thus, even if a Gaussian is centered on the theoretical position of the absorption ($\mu = 494$ nm), Gaussians with small amplitudes are also selected in this area.

3) SWIR absorptions: kaolinite doublet and secondary absorptions: Fig. 8 shows the results for kaolinite. We focus on the absorption doublet at 2162 and 2206 nm and the secondary absorptions at 2312 and 2380 nm. The estimated continuum is removed for improved readability.

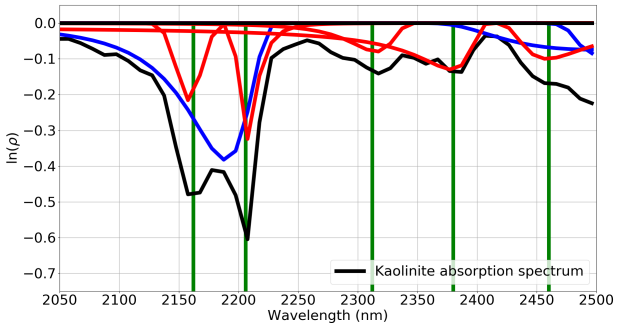


Fig. 8: SWIR kaolinite absorption spectrum (continuum removed, black), diagnostic and secondary Gaussians (red), other Gaussians (blue). The true positions of the diagnostic and secondary estimated Gaussians are shown as vertical lines.

The main kaolinite doublet is fitted by three EGO Gaussians. The broad Gaussian at 2189 nm (in blue in Fig. 8) models the shape of the absorption while the other two (in red), narrow and deep, are located close to the theoretical positions of the doublet ($\delta_\mu < 4$ nm). This result is explained by the fact that NNOMP first selects the broad Gaussian to capture the main trends of the spectrum. Nevertheless, this solution, with three EGO Gaussians, is well-suited to a mineralogical interpretation, based on the positions of the doublet.

The secondary absorptions at 2312 and 2380 nm are fitted by two EGO Gaussians. Their positions are close to the theoretical ones ($\delta_\mu < 3$ nm). The absorptions being close, their amplitudes and shapes are difficult to separate. Thus, the EGO Gaussian centered at 2380 nm is asymmetric ($k = 0.43$) which reduces the amplitude of the one at 2312 nm.

The spectral range at the edge of the spectrum (for the highest wavelengths) is modeled by three EGO Gaussians. The Gaussian centered at 2458 nm corresponds to an absorption embedded in the spectrum, which is not characteristic of kaolinite. The wide Gaussian, centered at 2503 nm, compensates for a difference between the continuum and the spectrum. The Gaussian centered at 2495 nm fits a small trend on the edge of the spectrum. These three Gaussians are not useful for identification purpose, since they either fit a non diagnostic absorption or compensate artifacts on the edge of the spectrum.

4) VNIR and SWIR absorptions: nontronite continuum and absorptions: Fig. 9 shows the results for nontronite. We focus on the continuum shape, the diagnostic absorptions at 660, 960 and 2283 nm and the secondary absorption at 2378 nm.

The estimated continuum lays above the spectrum and fits the overall shape. Due to the presence of broad absorptions

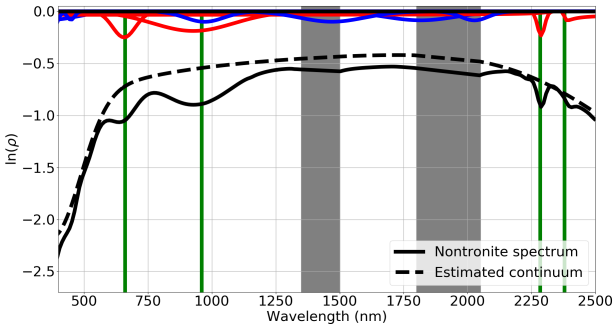


Fig. 9: Nontronite spectrum (black), estimated continuum (black-dashed), diagnostic and secondary Gaussians (red), other Gaussians (blue). The true positions of the diagnostic and secondary estimated Gaussians are shown as vertical lines.

in the VNIR, several artifacts are created. Three small and narrow EGO Gaussians are selected in the VNIR edge of the spectrum. Their positions being close to 450 nm, they cannot be interpreted as absorptions. The artifact between 600 and 1250 nm impacts the amplitude of the absorptions at 660 and 960 nm. It is compensated by a smaller EGO Gaussian at 971 nm. The broad artifact between 1250 and 2100 nm is fitted by three EGO Gaussians centered within the spectral masks (at 1470, 1816 and 2027 nm).

Nevertheless, all the absorptions of interest are retrieved and fitted by unique EGO Gaussians, highlighted in red in Fig. 9. Specifically, one can distinguish two Gaussians in the VNIR, the Gaussian of largest width being slightly shifted from the theoretical position at 960 nm. Two other Gaussians are found in the SWIR part of the spectrum. They correspond to the diagnostic absorption at 2283 nm and the secondary one at 2378 nm. Positions and widths are consistent with theoretical knowledge. However, due to a gap between the continuum and the spectrum, the shape of the secondary absorption is modified. This results in a strong asymmetry of the Gaussian at 2378 nm ($k = -0.73$).

5) Spectral mask: gypsum absorptions: Fig. 10 synthesizes the results for gypsum. The estimated absorptions located outside and within the spectral mask are shown in Fig. 10a and Fig. 10b, respectively. We focus on the diagnostic absorption at 1750 nm and the secondary absorptions at 1538 and 2215 nm, shown with vertical lines in Fig. 10a. Note that the absorptions at 1538 and 1750 nm are close to the spectral masks. Therefore, they are more difficult to retrieve.

In Fig. 10a, the diagnostic absorption at 1750 nm is fitted by an EGO Gaussian whose position and shape are slightly impacted by missing data within the spectral mask ($\delta_\mu = 2$ nm). Regarding the secondary absorption at 1538 nm, two Gaussians are retrieved. The first, centered at 1555 nm (in blue), is wide and deep and fits the right edge of the absorption. The second, of lower amplitude and width, gives the exact position of the absorption (in red). At last, the broad absorption at 2215 nm is fitted by a broad EGO Gaussian giving the shape of the absorption and a smaller one, in amplitude and width, positioned at the center of the absorption ($\mu = 2218$ nm). Note that an absorption that is not characteristic of gypsum,

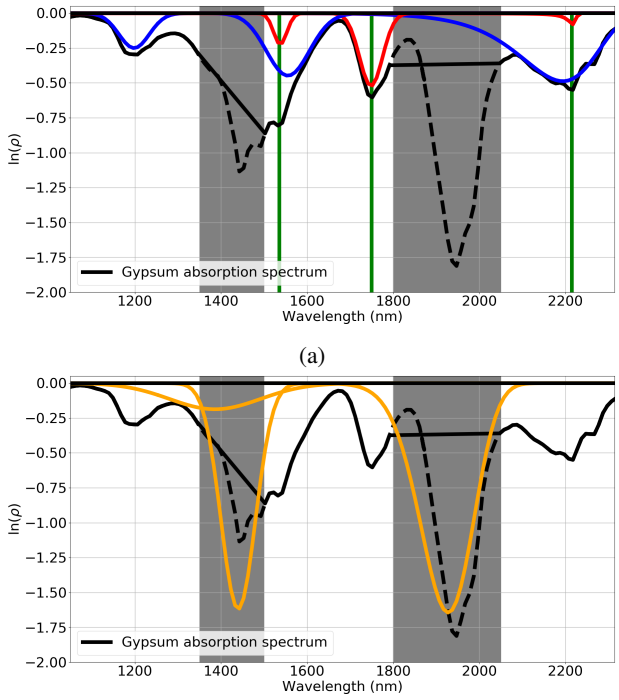


Fig. 10: Gypsum absorption data spectrum including missing data in the spectral masks (plain black; continuum has been removed. For information, the dashed curve indicates the full absorption spectrum). (a). Diagnostic and secondary Gaussians (red), and other Gaussians (blue). The positions of true Gaussians are shown as vertical lines. (b) Estimated Gaussians located in the spectral mask (orange).

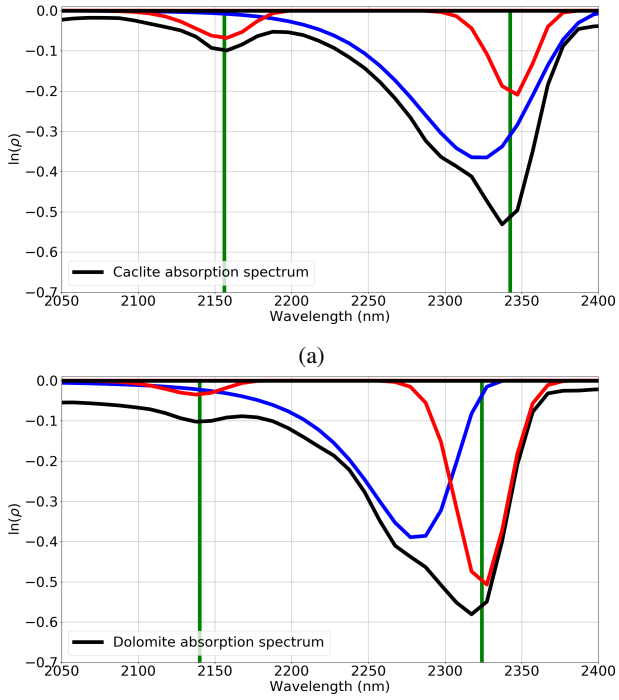


Fig. 11: Calcite (a) and dolomite (b) absorption spectrum (black), diagnostic and secondary absorptions (red) and other absorption (blue). The positions of the true diagnostic and secondary estimated Gaussians are shown as vertical lines.

is retrieved at 1198 nm.

In Fig. 10b, the edges of each spectral mask centered around 1400 and 1900 nm, respectively, are fitted by unique EGO Gaussians whose positions are close to the center of the masks. Their shape and position are consistent with the theoretical absorptions in the spectral mask, shown in dashed curve. Note that a third (broad) Gaussian is retrieved at 1386 nm, which models the left edge of this spectral mask.

This example shows that it is possible to retrieve absorptions centered close to the spectral masks despite missing data. The NNOMP algorithm indeed selects Gaussians centered in the spectral mask in order to adjust the edges of the spectrum. In the SWIR, we have to distinguish Gaussians, which fit the trends of the large absorptions from the narrow one, yielding the absorption center.

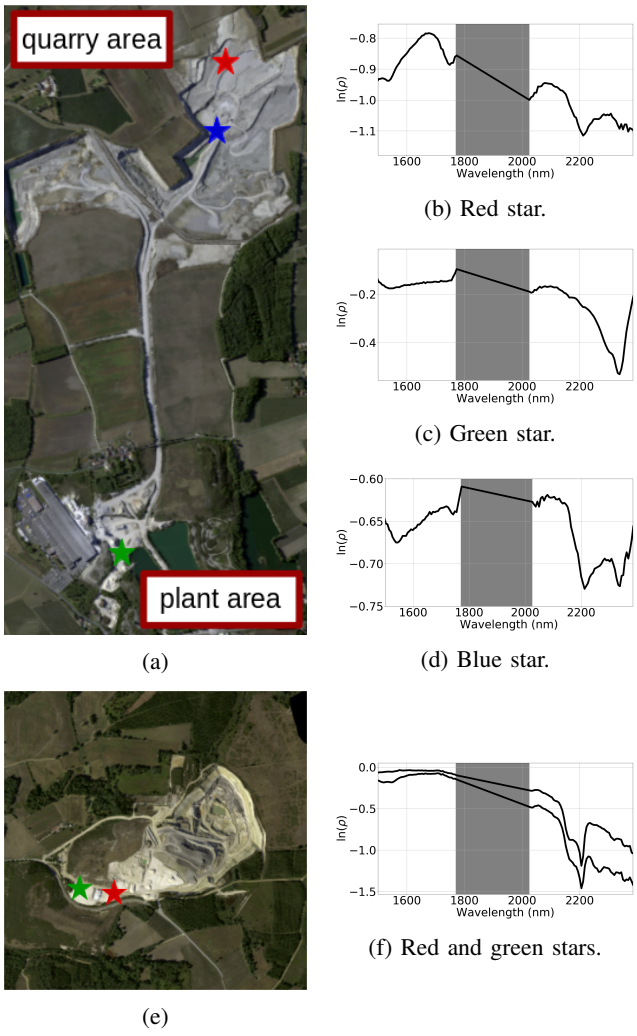
6) Close absorption discrimination: calcite and dolomite SWIR absorptions: Fig. 11 gathers the results for calcite and dolomite spectra, whose shapes are similar. We focus on their close absorptions (diagnostic is separated by 18.5 nm and secondary by 16 nm). The AVIRIS spectral resolution being around 10 nm, it is theoretically possible to distinguish the two minerals based on their absorption positions. To improve readability, the estimated continuum is removed from the spectrum of Fig. 11.

The absorptions of both spectra are deconvolved similarly. Two EGO Gaussians fit the main diagnostic absorption and a separate one is related to the secondary absorption. Similar to kaolinite, the blue Gaussian, broad and asymmetric ($\sigma = 37.5$ nm and $k = 0.13$ for calcite, $\sigma = 27.8$ nm and $k = 0.22$ for dolomite) models the main trend of the diagnostic absorption. The narrower Gaussian in red ($\sigma = 13.7$ nm for calcite and $\sigma = 16.5$ nm for dolomite) is located on the theoretical center of the absorption ($\delta_{mu} < 2$ nm). Its amplitude varies depending on how the Gaussian shape models the trend. Thus, for calcite, the amplitude of the Gaussian is lower than for dolomite. The secondary absorption is found close to the theoretical position ($\delta_{mu} < 2$ nm) with a visually coherent shape. However, for dolomite, its amplitude is attenuated by the presence of a wide Gaussian (not represented on Fig. 11b), centered before 2050 nm and resulting from a continuum estimation artifact.

This example shows that the diagnostic and secondary absorptions of calcite and dolomite are retrieved by the method. They can be used to distinguish calcite from dolomite.

C. Results on hyperspectral images

1) Campaign and data description: The method is now applied to spectra from hyperspectral images (Fig. 12) acquired during a dedicated campaign over two quarries in France, with HySpex cameras (https://www.hyspex.com). Three minerals of interest, with specific spectral signatures, can be observed: gypsum, carbonates (Cherves-Richemont) and kaolinite (Chevanceaux). They entail different physical and chemical



(a)

(b) Red star.

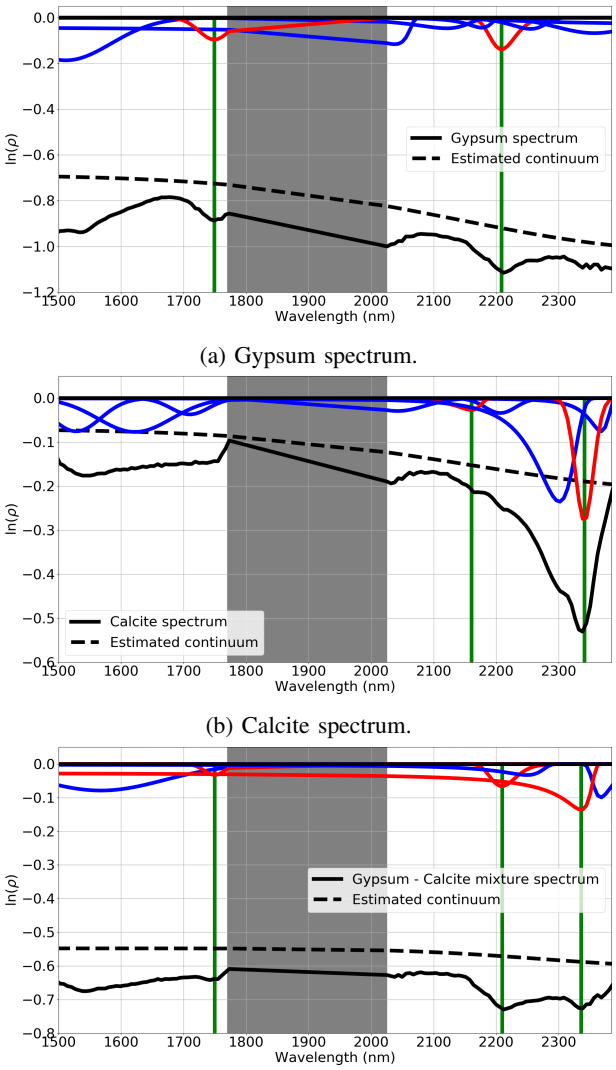
(c) Green star.

(d) Blue star.

(e)

(f) Red and green stars.

Fig. 12: Hyperspectral images acquired over (a) Cherves-Richemont and (e) Chevanceaux quarries. (b) Gypsum spectrum on the quarry layer. (c) Calcite spectrum from piles used as aggregates. (d) Gypsum-calcite mixture. (f) Kaolinite spectra from two piles.



(a) Gypsum spectrum.

(b) Calcite spectrum.

(c) Gypsum-calcite mixture

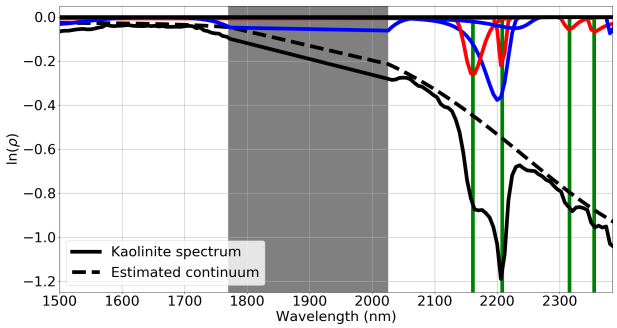
Fig. 13: Deconvolution of gypsum and calcite spectra, and a mixture of both. The true positions of the diagnostic and secondary absorptions are shown as vertical lines.

characteristics and are used in the production of plasterboard or as aggregates (Cherves-Richemont) and refractory ceramics (Chevanceaux). Images were acquired in September 2019, with a 0.5 m and a 1.0 m spatial resolution for the VNIR and SWIR, respectively. The number of bands were respectively 160 and 162 for the VNIR and SWIR with a ~ 4 nm and ~ 7 nm spectral resolution. In this study, only SWIR images are used. Images were atmospherically corrected using ATCOR4 [49]. To improve the SNR, reflectance images were spatially downsampled to 5 m. Two spectral masks were applied around 1400 and 1900 nm to avoid effects of saturated atmospheric water vapor absorptions. Samples were collected after image acquisition in order to create a spectral database gathering the possible spectral signatures in the quarries. An ASD FieldSpec®FR3 was used for laboratory measurements. The goals are to evaluate the performance of the method on spectra from hyperspectral images and to describe the potential of the EGO model for physical interpretation of various mineral reflectance spectra.

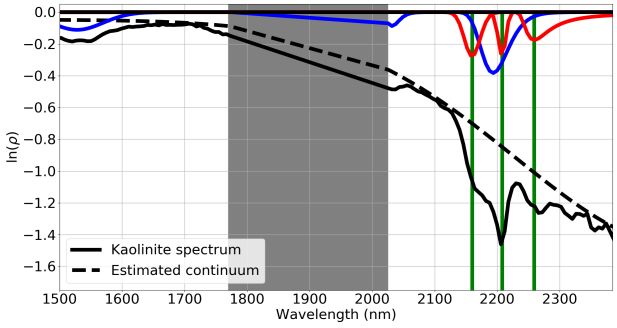
Hereafter, we analyze a mixture of gypsum and carbonate (here calcite) and spectra of kaolinite. To apply the method on SWIR spectra, we removed c_1 and G_{uv} from the continuum model (3), and generated a NNOMP dictionary, section III-C4, without EGO Gaussians centered shorter than 1500 nm. The noise covariance $\mathbf{\Sigma}$ is estimated using HYSIME [30].

2) Spectral mixture of gypsum and carbonates: In the Cherves-Richemont image, gypsum and calcite can be easily identified from their reflectance spectra (Fig. 12b and 12c). However, for some areas, especially near the roads and the plants, gypsum and calcite can be mixed. This results in a spectrum including all the diagnostic and secondary absorptions of gypsum and calcite (Fig. 12d). We thus apply the method to three spectra; see Fig. 13.

The results on gypsum and calcite can be easily compared to those on laboratory data (sections V-B5 and V-B6, respec-



(a)



(b)

Fig. 14: Deconvolution results on kaolinite spectra from two piles. The true positions of the diagnostic and secondary estimated Gaussians are shown as vertical lines.

tively). In Fig. 13a, the gypsum absorption at 1750 nm is retrieved at $\mu = 1749$ nm with an amplitude of 0.16 (*i.e.*, the sum of all estimated EGO Gaussians at 1749 nm). The broad absorption at 2215 nm is retrieved at $\mu = 2208$ nm using three EGO Gaussians. Also, the continuum parameter c_0 is equal to 0.67. In Fig. 13b, as expected, the calcite diagnostic absorption is fitted by two EGO Gaussians at ~ 2269 nm, and one for the theoretical position ($\mu = 2341$ nm). The secondary absorption at 2156 nm is retrieved at $\mu = 2160$ nm. The continuum parameter c_0 is equal to 0.07.

The diagnostic absorptions of gypsum and calcite and the secondary absorption of gypsum at 2215 nm are retrieved in the mixture (Fig. 13c). First, the absorption at 1750 nm is retrieved at $\mu = 1749$ nm with an amplitude of 0.08 (*i.e.*, the sum of all estimated EGO Gaussians at $\mu = 1749$ nm), which is half the one of the pure gypsum. Then, the absorption at 2215 nm is also retrieved at $\mu = 2209$ nm. The calcite absorption at 2340 nm does not match the shape of pure calcite. Thus, only one asymmetric EGO Gaussian is retrieved at 2235 nm. Finally, the continuum parameter c_0 is equal to 0.54, which is between the values for gypsum and calcite.

Based on the estimated EGO Gaussian positions and amplitudes, one can identify the spectrum as a gypsum-calcite mixture (positions) and possibly evaluate the relative concentration of both (amplitudes and c_0). This illustrates the high potential of spectral deconvolution for mineralogical interpretation.

3) Deconvolution of kaolinite spectra: On the Chevanceaux image, kaolinite is easily identifiable from its reflectance spectra (Fig. 12f). However, depending on the pile, the kaolinite doublet is more or less well formed, which may be related to crystallinity [1]. Results for each spectrum of Fig. 12f are shown in Fig. 14. They can be easily compared to those of laboratory data in paragraph V-B3. Note that an absorption at 2260 nm, probably due to the presence of gibbsite ($AlOOH$), is visible in the spectrum of Fig. 14b.

For both spectra, the absorption doublet is fitted by a broad EGO Gaussian (in blue) capturing the main shape and two narrow ones (in red). The broad Gaussian is positioned at 2201 and 2193 nm, respectively, with a width equal to 17.6 and 21.5 nm, respectively. The other two EGO Gaussians modeling the absorption doublet, are shifted from less than 1 nm from their true positions. Their amplitudes are equal to 0.42 and 0.63. For the spectrum of Fig. 14a, the small absorptions at 2315 and 2355 nm are well retrieved. On the contrary, in Fig. 14b, the strong asymmetry of the gibbsite absorption at 2259 nm does not make it possible to detect the small absorptions.

Similar to the gypsum-calcite mixture from the Cherves-Richemond image in paragraph V-C2, the diagnostic absorptions of gibbsite and kaolinite are retrieved, which allows one to identify these two minerals. Also, the kaolinite absorption doublet is fitted similarly for both spectra, so studying the evolution of its shape over the piles is possible.

VI. Conclusion

We proposed a spectral deconvolution procedure, automatic and adapted to the various absorption shapes in the VNIR and SWIR. This method extracts the main features (continuum and absorptions) from an arbitrary spectrum without prior knowledge on the mineralogical context. It relies on the EGO physical model. It was first validated on synthetic spectra representative of various situations that can be encountered in spectral signatures of minerals: broad and shallow absorptions such as those related to iron in the VNIR, narrow and sharp, close absorptions like the kaolinite doublet, absorptions located close to the atmospheric water vapor absorptions, asymmetric absorptions, *etc*. The number of absorptions and their respective parameters were well retrieved. For example, errors are less than 20 nm and 5 nm for respectively VNIR and SWIR absorption positions. Also, asymmetric absorptions are fitted. The method was then applied to well characterized laboratory spectra. All the main features were retrieved and a physical interpretation is thus possible based on the estimated parameters. Finally, the method was applied to spectra from hyperspectral images acquired during a dedicated campaign. The method was able to handle mixtures of gypsum and calcite, and the kaolinite absorption doublet was well retrieved.

The method is flexible since other models or constraints could be taken into account. For example, a better understanding of mineral spectroscopy may lead to improve the EGO model. Similarly, the constraints in the continuum pre-estimation step, related to the continuum physics can be adjusted. Also, the NNOMP dictionary can be created for specific absorption shapes with more than four parameters and could take into account prior information on the number of absorptions and their respective positions and shapes.

This study also demonstrates the potential of the method to obtain parameters allowing the identification and characteriza-

tion of a mineral from its reflectance spectrum. An automatic identification procedure, based on a fuzzy logic system, has been proposed by the authors [50]. It uses as input the EGO parameters and takes into account uncertainties that can be computed after the joint optimization step. The AGM needs to be tested on more images from various hyperspectral sensors and related to diverse geological environments. This would allow one to assess the parameter estimation uncertainties as a function of instrument performances, and then to produce maps of the various parameters.

ACKNOWLEDGMENT

The authors acknowledge the LPG Nantes and GEOFIT EXPERT for the acquisition of the HySpex images. Imerys and Garandeau firms are thanked for their collaboration.

REFERENCES

[1] R. N. Clark, *Manual of remote sensing, volume 3, remote sensing for the Earth sciences*. John Wiley and Sons, New York, 1999, ch. 1: Spectroscopy of rocks and minerals, and principles of spectroscopy, pp. 3–58.

[2] R. Marion and V. Carrère, "Mineral mapping using the Automatized Gaussian Model (AGM) - Application to two industrial French sites at Gardanne and Thann," *Remote Sensing*, vol. 10, no. 1, 2018.

[3] R. Green, M. Eastwood, C. Sarture, T. Chrien, M. Aronsson, B. Chippendale, J. Faust, B. Pavri, C. Chovit, M. Solis, M. Olah, and O. Williams, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, 1998.

[4] L. Hamlin, R. O. Green, P. Mouroulis, M. Eastwood, D. Wilson, M. Dudik, and C. Paine, "Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments," in *Aerospace Conference*, Mar. 2011, pp. 1–7.

[5] R. Loizzo, R. Guarini, F. Longo, T. Scopa, R. Formaro, C. Facchinetti, and G. Varacalli, "Prisma: The Italian hyperspectral mission," in *IEEE Int. Geosc. Rem. Sens. Symp.*, Jul. 2018, pp. 175–178.

[6] L. Guanter, H. Kaufmann, K. Segl, S. Foerster, C. Rogass, S. Chabrillat, T. Kuester, A. Hollstein, G. Rossner, C. Chlebek, C. Straif, S. Fischer, S. Schrader, T. Storch, U. Heiden, A. Mueller, M. Bachmann, H. Mühle, R. Müller, M. Habermeyer, A. Ohndorf, J. Hill, H. Buddenbaum, P. Hostert, S. Van der Linden, P. J. Leitao, A. Rabe, R. Doerffer, H. Krasemann, H. Xi, W. Mauser, T. Hank, M. Locherer, M. Rast, K. Staenz, and B. Sang, "The EnMAP spaceborne imaging spectroscopy mission for Earth observation," *Remote Sensing*, vol. 7, no. 7, pp. 8830–8857, 2015.

[7] S. Asadzadeh and C. R. de Souza Filho, "A review on spectral processing methods for geological remote sensing," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 47, pp. 69 – 90, 2016.

[8] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, 2012.

[9] F. van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery," *Int. J. Appl. Earth Obs.*, vol. 8, no. 1, pp. 3 – 17, 2006.

[10] R. Kokaly, R. N. Clark, G. Swayze, K. E. Livo, T. Hoefen, N. Pearson, R. Wise, W. Benzel, H. Lowers, and R. L. Driscoll, "USGS Spectral Library Version 7," U.S. Geological Survey, Report, Apr. 2017.

[11] S. Chabrillat, A. Eisele, S. Guillaso, A. Eisele, C. Rogaß, E. Ben-Dor, and H. Kaufmann, "HYSOMA: An easy-to-use software interface for soil mapping applications of hyperspectral imagery," *Proceedings of the 7th EARSeL SIG Imaging Spectroscopy Workshop*, 2011.

[12] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. M. R. R., and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res. Planets*, vol. 108, no. E12, p. 5131, 2003.

[13] C. Hecker, F. J. A. van Ruitenbeek, H. M. A. van der Werff, W. H. Bakker, R. D. Hewson, and F. D. van der Meer, "Spectral absorption feature analysis for finding ore: A tutorial on using the method in geological remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 51–71, Jun. 2019.

[14] V. Kopačková and L. Koucká, "Mineral mapping based on automatic detection of multiple absorption features," *EARSeL eProceedings*, vol. 13, no. S1, pp. 95–99, 2014.

[15] J. M. Sunshine, C. M. Pieters, and S. F. Pratt, "Deconvolution of mineral absorption bands: An improved approach," *J. Geophys. Res. Solid Earth*, vol. 95, no. B5, pp. 6955–6966, 1990.

[16] L. Pompilio, G. Pedrazzi, M. Sgavetti, E. A. Cloutis, M. A. Craig, and T. L. Roush, "Exponential Gaussian approach for spectral modeling: The EGO algorithm I. Band saturation," *Icarus*, vol. 201, no. 2, pp. 781 – 794, 2009.

[17] L. Pompilio, G. Pedrazzi, E. A. Cloutis, M. A. Craig, and T. L. Roush, "Exponential Gaussian approach for spectral modelling: The EGO algorithm II. Band asymmetry," *Icarus*, vol. 208, no. 2, pp. 811 – 823, 2010.

[18] H. D. Makarewicz, M. Parente, and J. L. Bishop, "Deconvolution of VNIR spectra using modified Gaussian modeling (MGM) with automatic parameter initialization (API) applied to CRISM," in *First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Aug. 2009, pp. 1–5.

[19] H. Clénet, P. Pinet, Y. Daydou, F. Heuripeau, C. Rosemberg, D. Baratoux, and S. Chevrel, "A new systematic approach using the Modified Gaussian Model: Insight for the characterization of chemical composition of olivines, pyroxenes and olivine–pyroxene mixtures," *Icarus*, vol. 213, no. 1, pp. 404 – 422, 2011.

[20] C. Verpoorter, V. Carrère, and J.-P. Combe, "Visible, near-infrared spectrometry for simultaneous assessment of geophysical sediment properties (water and grain size) using the Spectral Derivative–Modified Gaussian Model," *J. Geophys. Res. Earth Surf.*, vol. 119, no. 10, pp. 2098–2122, 2014.

[21] M. Brossard, R. Marion, and V. Carrère, "Deconvolution of SWIR reflectance spectra for automatic mineral identification in hyperspectral imaging," *Remote Sens. Lett.*, vol. 7, no. 6, pp. 581–590, 2016.

[22] J. M. Sunshine and C. M. Pieters, "Estimating modal abundances from the spectra of natural and laboratory pyroxene mixtures using the Modified Gaussian Model," *J. Geophys. Res. Planets*, vol. 98, no. E5, pp. 9075–9087, 1993.

[23] A. J. Brown, "Spectral curve fitting for automatic hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1601–1608, 2006.

[24] V. L. Mulder, M. Plötze, S. de Bruin, M. E. Schaepman, C. Mavris, R. F. Kokaly, and M. Egli, "Quantifying mineral abundances of complex mixtures by coupling spectral deconvolution of SWIR spectra (2.1–2.4 μm) and regression tree analysis," *Geoderma*, vol. 207-208, pp. 279 – 290, 2013.

[25] R. N. Clark and T. L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *J. Geophys. Res. Solid Earth*, vol. 89, no. B7, pp. 6329–6340, 1984.

[26] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures." *J. Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[27] M. L. Whiting, L. Li, and S. L. Ustin, "Predicting water content using gaussian model on soil spectra," *Remote Sens. Environ.*, vol. 89, no. 4, pp. 535 – 552, 2004.

[28] P. Du, W. A. Kibbe, and S. M. Lin, "Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching," *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, Jul. 2006.

[29] B. Rivard, J. Feng, A. Gallie, and A. Sanchez-Azofeifa, "Continuous wavelets for the improved use of spectral libraries and hyperspectral data," *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2850 – 2862, 2008.

[30] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[31] M. J. D. Powell, "Direct search algorithms for optimization calculations," *Acta Numerica*, vol. 7, p. 287–336, 1998.

[32] P. Stoica and Y. Selén, "Model-order selection: A review of information criterion rules," *IEEE Sig. Proc. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.

[33] A. M. Bruckstein, M. Elad, and M. Zibulevsky, "On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4813–4820, 2008.

[34] T. T. Nguyen, J. Idier, C. Soussen, and E. Djermoune, "Non-negative orthogonal greedy algorithms," *IEEE Trans. Signal Process.*, vol. 67, no. 21, pp. 5643–5658, Nov. 2019.

[35] C. Soussen, J. Idier, J. Duan, and D. Brie, "Homotopy based algorithms for ℓ_0-regularized least-squares," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3301–3316, Jul. 2015.

[36] F. De Ridder, R. Pintelon, J. Schoukens, and D. P. Gillikin, "Modified AIC and MDL model selection criteria for short data records," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 1, pp. 144–150, Feb. 2005.
[37] M. A. Branch, T. F. Coleman, and Y. Li, "A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems," *SIAM J. Sci. Comput.*, vol. 21, no. 1, pp. 1–23, 1999.
[38] S. Pontual, N. Merry, and P. Gamson, "GMEX: Guides for Mineral EXploration: Spectral interpretation field manual," *AusSpec International Ltd.: Queenstown, New Zealand*, vol. 1, p. 191, 2010.
[39] G. R. Hunt, J. W. Salisbury, and C. Lenhoff, "Visible and near-infrared spectra of minerals ans rocks : III. Oxides and hydroxides," *Mod. Geol.*, vol. 2, pp. 195–205, 1971.
[40] G. R. Hunt and J. W. Salisbury, "Visible and near-infrared spectra of minerals ans rocks : I. Silicate minerals," *Mod. Geol.*, vol. 1, pp. 283–300, 1970.
[41] G. R. Hunt, J. W. Salisbury, and C. Lenhoff, "Visible and near-infrared spectra of minerals ans rocks : VI. Additional silicates," *Mod. Geol.*, vol. 4, pp. 85–106, 1973.
[42] J. L. Bishop, M. D. Lane, M. D. Dyar, and A. J. Brown, "Reflectance and emission spectroscopy study of four groups of phyllosilicates: Smectites, kaolinite-serpentines, chlorites and micas," *Clay Minerals*, vol. 43, no. 1, pp. 35–54, Mar. 2008.
[43] J. Flahaut, M. Martinot, J. L. Bishop, G. R. Davies, and N. J. Potts, "Remote sensing and in situ mineralogic survey of the Chilean salars: An analog to Mars evaporate deposits?" *Icarus*, vol. 282, pp. 152 – 173, 2017.
[44] G. R. Hunt, J. W. Salisbury, and C. Lenhoff, "Visible and near-infrared spectra of minerals ans rocks : IV. Sulphides and sulphates," *Mod. Geol.*, vol. 3, pp. 1–14, 1971.
[45] E. A. Cloutis, F. C. Hawthorne, S. A. Mertzman, K. Krenn, M. A. Craig, D. Marcino, M. Methot, J. Strong, J. F. Mustard, D. L. Blaney, J. F. Bell, and F. Vilas, "Detection and discrimination of sulfate minerals using reflectance spectroscopy," *Icarus*, vol. 184, no. 1, pp. 121 – 157, 2006.
[46] G. R. Hunt and J. W. Salisbury, "Visible and near-infrared spectra of minerals ans rocks : II. Carbonates," *Mod. Geol.*, vol. 2, pp. 23–30, 1971.
[47] S. J. Gaffey, "Spectral reflectance of carbonates minerals in the visible and near-infrared $(0.35 - 2.55$ microns$)$: Calcite, aragonite and dolomite," *Am. Minerals*, vol. 712, pp. 151–162, 1986.
[48] E. A. Cloutis, S. E. Grasby, W. M. Last, R. Léveillé, G. R. Osinski, and B. L. Sherriff, "Spectral reflectance properties of carbonates from terrestrial analogue environments: Implications for Mars," *Planetary and Space Science*, vol. 58, no. 4, pp. 522 – 537, 2010.
[49] R. Richter and D. Schlapfer, "Atmospheric/topographic correction for airborne imagery: ATCOR-4 user Guide," *DLR IB*, pp. 565–02, 2012.
[50] R. Rialland, V. Carrère, R. Marion, and C. Soussen, "Identification of minerals from hyperspectral imaging based on a fuzzy logic approach," in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXVII*, M. Velez-Reyes and D. W. Messinger, Eds., vol. 11727. SPIE, 2021, pp. 108 – 119.