RETRIEVAL OF PHYSICAL PROPERTIES OF MUDFLAT SEDIMENTS FROM HYPERSPECTRAL DATA USING THE MODIFIED GAUSSIAN MODEL AND SPECTRAL CURVE FITTING

Charles Verpoorter\textsuperscript{1}, Véronique Carrère\textsuperscript{1}, Marc Robin\textsuperscript{2}

1. Laboratoire de Planétologie et Géodynamique UMR-CNRS 6112, BP 92 208, F -44322 Charles.Verpoorter@univ-nantes.fr
2. Géolittomer, LETG UMR-CNRS 6554, BP 81 227, -F -44312

ABSTRACT

The objective of this work is to retrieve parameters relevant to the understanding of sediment dynamics in a bay, Bourgneuf Bay (France) from hyperspectral remote sensing data. The erodability of mudflats is strongly influenced by the sediments bio-physical characteristics. These parameters include sediment grain size, moisture content and mineralogy as well as presence or absence of biofilm, all having an influence on sediment cohesiveness. The surface spectral signature gives access to mineralogy through specific absorption features. Sediment grain size influences the scattering / absorption relationship which translates into band shape, average surface reflectance and the general shape of the spectrum also called the continuum. Moisture content also influences the shape of the continuum as well as the strength of the specific liquid water absorption features observable in the NIR. Finally chlorophyll-a concentration derived from the strength of the absorption at 675 nm, is a good proxy for biofilm biomass. In order to access all these characteristics, we used the Modified Gaussian Model [MGM] to deconvolve the reflectance spectra into constituent absorptions and a continuum curve. This inversion process was semi-automatic. In this paper, an automated method for hyperspectral data is proposed, using as a first step spectroscopic derivative analysis to calculate an initial solution for the MGM input parameters.

Keywords: Modified Gaussian Model, derivative analysis, ROSIS, sediments parameters.

INTRODUCTION

Mineral composition, mineral mixing, grain size and moisture content are important factors influencing the spectral signature of sediment. They are also important parameters for the understanding of sediment dynamics in the intertidal zone because they control sediment cohesiveness and therefore erodability. An additional factor that also needs to be taken into account is the presence or absence of biofilm. The aim of this work is to develop an algorithm as automated as possible to get access to the physical properties of mudflats sediments from hyperspectral data and therefore map their spatial distribution.

Minerals exhibit a number of distinct absorptions reflecting their composition and type. These absorptions vary in position, relative strength and shape. Additionally, scattering process contributes significantly to the overall shape of the reflectance spectrum or continuum. The continuum envelope of a spectrum is the overall reflected light color of an object relative to incident light. The continuum is related to physical properties of the surface (e.g. grain size, roughness, moisture content, local slope, etc.). It includes the contribution of many radiative processes that are not explained by absorption bands. Therefore, the shape of the continuum may change from place to place for a homogeneous surface composition. In order to establish relationships between spectral information and physical properties of mudflat sediments, it is necessary to isolate absorptions from the overall shape or continuum. Clark and Roush (1984), (i), have developed a simple and easy-to-use approach (Figure 1a) that is implemented in the ENVI software package. However, this approach prevents from properly modeling absorption features that may exist at the edge of the wavelength range considered, as, for example, for plant pigments in the VIS. Our approach is thus based on the Modified Gaussian Model [MGM] (ii, iii), a semi-automatic process

...
that deconvolves reflectance spectra into a continuum and absorption bands and allows to take this particular case into account. Additionally, MGM has proven to be one of the most powerful tools for deconvolution of reflectance spectrum of binary or ternary mixtures as well as overlapping absorption features (Figure 1a). Absorption features derived from MGM deconvolution are assimilated to Gaussian shapes. These absorption shapes can be defined by their strength, centers and Full-Width at Half-Maximum [FWHM]. Absorption strength can in turn be related, for example, to relative mineral abundances in mixtures while central position is indicative of composition and crystal structure.

However, the MGM deconvolution is a semi-automatic process and it needs an initial set of parameters, representing a priori constraints. Due to the large size of hyperspectral datasets, more robust and stable automated methods are desirable. Moreover, the MGM results are highly dependant on this initial set of parameters. We thus coupled the MGM analysis to a derivative analysis to reduce a priori knowledge for the initial set of parameters. In order to validate the methodology, we applied this process first to spectra from the USGS spectral library (including mixtures) as well as to field spectra, then to hyperspectral ROSIS data acquired over Bourgneuf bay, SW France. This method allows to extract spectral characteristics from the image and should give information about parameters characterizing mudflat sediments.

Figure 1: Example of Kaolinite-Smectite mixture (KLF 508 85%K) spectrum from USGS spectral library. a) Continuum removal using the straight line approach; solid line, original spectrum (bottom) and after continuum removal (top), dashed line, continuum. b) Continuum removal using the MGM approach; solid line, original spectrum (bottom) and after continuum removal (top), dashed line, continuum; dashed line (top), Gaussian functions.
METHODOLOGY

Principle of the Modified Gaussian deconvolution

The natural logarithm of a reflectance spectrum can be deconvolved as a sum of \( n \) absorption bands superimposed on a baseline background so called continuum (Figure 1b). The algorithm applies a non-linear least square inversion based on the stochastic technique of Tarantola and Valette (iv). Each absorption band is expressed as a Modified Gaussian in wavelength whereas the continuum is a straight line in wavenumber. The latter can be described by a slope and an offset. The MGM formula can be expressed as (ii):

\[
\ln(R_k) = (a\lambda_k^{-1} + b) + \sum_{i=1}^{n} s_i \exp \left[ -\frac{(\lambda_k - \mu_i)^2}{2\sigma_i^2} \right] \quad (1)
\]

Where \( R_k \) is the measured reflectance at a given wavelength \( \lambda_k \) and band parameters \( \sigma, \mu, s \), are the width (= standard deviation), central wavelength and strength. The Full-Width at Half-Maximum [FWHM] is calculated by:

\[
FWHM = 2\sqrt{2 \ln 2} \sigma \cong 2.35484 \sigma \quad (2)
\]

The background continuum component \( C(\lambda_k) \) is expressed as:

\[
C(\lambda_k) = \frac{a}{\lambda_k} + b \quad (3)
\]

Where \( a \), the slope and \( b \), the intercept for the linear function in wavenumber, are supposed to be constant parameters inherent to the host constituent.

As the continuum is empirical, our first problem is to choose a reasonable representation that is consistent with the model run and in agreement with the scattering theory. We selected the continuum expressed in equation (3) because working in the energy (wavenumber) space is physically correct due to his connection with the energy of the electronic transition in the molecule of interest. A continuum with a negative slope, linear in energy, providing the best approximation of the blue continuum was used here.

Finally, the quality of the modeled spectrum is described by the Root Mean-Square [RMS]. The residual error which corresponds to the difference between the modeled \( R'(\lambda_i) \) and the measured natural logarithm reflectance spectrum \( R(\lambda_i) \) is minimized by the algorithm.

\[
RMS = \sqrt{\frac{1}{n} \sum_{i=1}^{n}(R(\lambda_i) - R'(\lambda_i))^2} \quad (4)
\]

Where \( \lambda \) is the wavelength, \( n \) the number of channels and \( i \) an index varying from 1 to \( n \).
The MGM deconvolution is a semi-automatic process and needs an initial set of parameters, representing *a priori* constraints. An example of input parameter file is given in Table 1.

**Table 1: Example of an initial set of parameters for MGM fitting: set of Gaussian and continuum parameters for deconvolution of a ROSIS reflectance spectrum (benthic diatoms).**

<table>
<thead>
<tr>
<th>Band</th>
<th>Central Position (nm)</th>
<th>FWHM (nm)</th>
<th>Biological interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>433±200</td>
<td>70±300</td>
<td>All pigments</td>
</tr>
<tr>
<td>2</td>
<td>493±200</td>
<td>76±300</td>
<td>β-carotene</td>
</tr>
<tr>
<td>3</td>
<td>500±200</td>
<td>60±300</td>
<td>Diadinoxanthin</td>
</tr>
<tr>
<td>4</td>
<td>550±200</td>
<td>60±300</td>
<td>Fucoxanthin</td>
</tr>
<tr>
<td>5</td>
<td>630±200</td>
<td>40±300</td>
<td>Chorophyll-c</td>
</tr>
<tr>
<td>6</td>
<td>674±200</td>
<td>14±300</td>
<td>Chorophyll-a</td>
</tr>
<tr>
<td>7</td>
<td>737±200</td>
<td>8±300</td>
<td>Cell structure</td>
</tr>
<tr>
<td>8</td>
<td>750±200</td>
<td>16±300</td>
<td>Cell structure</td>
</tr>
<tr>
<td>9</td>
<td>794±200</td>
<td>80±300</td>
<td>Cell structure</td>
</tr>
<tr>
<td>10</td>
<td>810±200</td>
<td>380±300</td>
<td>Improve MGM fit</td>
</tr>
</tbody>
</table>

As shown in Figure 2, results of the deconvolution are a set of Gaussian functions figuring absorptions and a continuum.

*Figure 2: Representative MGM fit for benthic diatoms (microphytobenthos). The horizontal dashed arrow indicates the FWHM at 630 nm [Chl-c]. The vertical dashed arrow indicates the strength of the absorption at 674 nm [Chl-a].*
Main steps for automatic retrieval of MGM input parameters

The main steps followed to minimize a priori knowledge necessary to select input parameters for the MGM are described in Figure 3.

Figure 3: Flow chart, representing the general methodology to automate the Modified Gaussian Model (example: Illite-Montmorillonite CM42 mixture from the USGS spectral library).
1. Noise removal:

If a spectrum shows a poor Signal/Noise ratio, additional minima can occur that have no spectroscopic significance. It is necessary to use a smoothing or filtering (Savitzky-Golay or Minimum Noise Fraction transform [MNF],...) method to reduce the effects of high frequency noise in the original spectral curve. MNF can show the same spectral attributes than the Savitzky-Golay method, a procedure based on a polynomial least-squares fit. In this study, we applied the polynomial smoothing-differentiation of Savitzky-Golay (v) to spectral libraries and the MNF transform to hyperspectral images. An appropriate smoothing level must be chosen to reduce noise taking into account that increasing smoothing may decrease significantly the spectral resolution and broaden band widths or attenuate band depth. The smoothing width should remain inferior to the width of the spectral features of interest.

2. Estimation of absorption features characteristics:

Once, when necessary, noise has been removed, the next step is to help establish automatically the input parameters related to absorption features: position and FWHM. A simple continuum removal is applied, following Clark and Roush (1984). The continuum is here represented by a tangent line to the local maxima (Figure 1a). Straight line segments are drawn between the defined high points of the spectrum. The continuum is then removed by dividing the reflectance spectrum by the continuum for each channel. Although this method is fast and reproducible, it is strongly sensitive to absorption bands, particularly at the edge of the spectra. It is applied here to rapidly access absorption parameters. The continuum removed spectrum is then used as a starting point for the next step, the spectral derivative analysis, that aims at estimating the initial set of parameters characterizing absorption features automatically from the spectra themselves.

3. Derivative analysis:

Derivative spectroscopy in the VNIR region is a useful technique to extract qualitative and quantitative information (vi, vii) from hyperspectral data sets in terrestrial environments. The derivative analysis tool allows to work with fine spectral resolution. This tool is commonly used, for example, to eliminate soil background interference from leaf absorption spectra. It allows to improve identification to extract information from overlapping bands features of pigments and chemical compositions (viii, ix, x, xi). Furthermore derivative techniques reduce the effects of turbidity in aquatic chlorophyll investigations (xii).

In this study, we also used derivative analysis to extract and isolate absorption features from reflectance spectra (xiii, xiv). First and second order derivatives have been used for band decomposition. The first derivative can be expressed as:

\[
\frac{dy}{d\lambda} = \frac{s(\lambda_j) - s(\lambda_i)}{\Delta\lambda} \quad (5)
\]

Where \(\Delta\lambda\) (bandwidth) is the separation between adjacent bands \(\Delta\lambda = \lambda_j - \lambda_i\), where \(\lambda_j > \lambda_i\) and the interval between band is constant (vi).

The second derivative can be derived from the first derivative. The \(n^{th}\) derivative is computed using the following equation:

\[
\frac{d^n y}{d\lambda^n} = \frac{d}{d\lambda} \left( \frac{d^{(n-1)} y}{d\lambda^{(n-1)}} \right) \quad (6)
\]
The ‘zero crossing’ for the first derivative helps localizing local maxima and minima in the spectrum. The second derivative gives access to asymmetry (high magnitude), minima and maxima (derivative sign) and spectrum slope (when equal to 0).

Figure 4 shows parameters (minima, maxima, inflection points) as detected from a continuum removed spectrum. Examples of parameters obtained through derivative analysis are presented in Table 2.

![Figure 4: Original spectrum of a Kaolinite-Smectite mixture (KLF 508 85%K) from the USGS spectral library (bottom); and after continuum removal (top); after fitting by spectral derivative computation (dashed line).](image)

### Table 2: Representative Kaolinite-Smectite mixture (KLF 508 85%K) from the USGS spectral library, derived from the spectral derivative shape identification.

<table>
<thead>
<tr>
<th>Absorption position (nm)</th>
<th>Reflectance</th>
<th>Band Depth</th>
<th>FWHM (nm)</th>
<th>Symmetry</th>
<th>Symmetry (Log)</th>
</tr>
</thead>
<tbody>
<tr>
<td>415</td>
<td>0.97993167</td>
<td>0.02006833</td>
<td>296</td>
<td>15.963734</td>
<td>1.203134483</td>
</tr>
<tr>
<td>563</td>
<td>0.87932766</td>
<td>0.12067234</td>
<td>149</td>
<td>0.86619573</td>
<td>-0.062383961</td>
</tr>
<tr>
<td>883</td>
<td>0.94924783</td>
<td>0.05075217</td>
<td>153</td>
<td>1.4956646</td>
<td>0.174834215</td>
</tr>
<tr>
<td>1408</td>
<td>0.80794577</td>
<td>0.19205423</td>
<td>55</td>
<td>0.85044521</td>
<td>-0.070353661</td>
</tr>
<tr>
<td>1914</td>
<td>0.69605532</td>
<td>0.30394468</td>
<td>80</td>
<td>2.4458731</td>
<td>0.388433921</td>
</tr>
<tr>
<td>2204</td>
<td>0.75154814</td>
<td>0.24845186</td>
<td>49</td>
<td>0.39990586</td>
<td>-0.398042232</td>
</tr>
<tr>
<td>2385</td>
<td>0.94375553</td>
<td>0.05624448</td>
<td>85</td>
<td>0.24782325</td>
<td>-0.605857952</td>
</tr>
<tr>
<td>2496</td>
<td>0.96259488</td>
<td>0.03740512</td>
<td>223</td>
<td>0.14400843</td>
<td>-0.841612084</td>
</tr>
</tbody>
</table>
Other interesting spectral parameters can also be derived from this technique. Band area is given by the sum of the $\text{Area}_{\text{Left}}$ and $\text{Area}_{\text{Right}}$ of the line through the center of the absorption feature. Asymmetry is the ratio of the $\text{Area}_{\text{Left}}$ to $\text{Area}_{\text{Right}}$ of the absorption centre. The use of the base 10 logarithm preserves the linearity when calculating asymmetry for bands with equal asymmetries to the left and right of the band minimum. The result is that symmetric band will have a value of 0, asymmetric band to the left will be negative and to the right, positive (Figure 5).

\[
\text{Asymmetry Factor (AF)} = \log \left( \frac{\text{Area}_{\text{L}}}{\text{Area}_{\text{R}}} \right)
\]

![Figure 5: Schematic diagram illustrating the absorption feature parameters (band position, strength and shape), definition of asymmetry.](image)

This spectral identification also presents the advantage of reducing the number of iterations of the MGM process. This step gives us a quick approximation of absorption band and saves a lot of time. Even if the absorption band is broad, the derivative analysis allows to find the true minimum with a high probability. A first set of initial MGM parameters is thus constructed. In parallel, the parameters for the MGM continuum have to be estimated.

4. Spectral continuum modeling:
A MGM continuum can simply be modeled for each spectrum according to equation (3) in the natural logarithm reflectance and wavenumber range. The continuum is modeled as a straight line in wavenumber, the slope of the spectrum is anchored to the maximum point of the spectrum. It is a first approximation of the MGM continuum parameters (slope and intercept).

5. MGM deconvolution:
The MGM iterative process is finally run, resulting in a set of Gaussian functions, a continuum and a modeled spectrum. If the RMS is minimal the spectrum modeled by the MGM can be accepted. If not, abnormal values are used to readjust input parameters: add new Gaussian functions for overlapping absorptions not detected originally, change FWHM, etc. The MGM deconvolution is run again with the adjusted set and the procedure is repeated until the global RMS reaches satisfactory levels.
RESULTS AND DISCUSSIONS

The performances of the MGM were first tested by fitting reflectance spectra from the USGS spectral library, including mineral mixtures, as well as laboratory and field spectral data acquired with an ASD FieldSpec 3 in Bourgneuf Bay, France. The instrument records a continuous spectrum in 2151 bands, ranging from 350 nm to 2500 nm with 1.1 nm spectral resolution. A spectralon reference panel was used to calibrate spectral data to downwelling solar radiation.

Figure 3 shows an example of such a test performed on a mineral mixture of Illite-Montmorillonite CM42 from the USGS spectral library. The resulting RMS for this case is 0.140. Although these first tests were extremely satisfactory, some limitations appeared while attempting to automate the procedure. The main problem concerns the resolution of overlapping individual bands in complicated spectrum (xv, xvi). After a first MGM fitting, some problems remain when modeling asymmetrical absorption features due to overlapping and superimposed bands. In order to estimate the wavelength position of overlapping bands and then find their parameters, we can use:

- Visual inspection based on knowledge and experience of the researcher.
- Graphical methods, for example the wavelength position of inflection point as well as the kurtosis (Magnitude of the 2nd derivative) or the asymmetry factor.
- Use the residual error [RMS] as described above to locate additional overlapping absorptions. This approach gives a better estimate of the mutual influence of the band as well as the effect of the continuum. It is a good indicator for wavelength position of overlapping bands but is very sensitive to noise. By sorting the RMS values in ascending order and isolating the five highest absolute values, a good estimate of the additional bands required for a better fit is obtained.

Finally, in the cases of spectra with complex shapes, the procedure fails to converge, generally because of noise level, difficulty to estimate a realistic continuum or the number of overlapping bands. In some cases, assuming a Modified Gaussian band shape function might not be a good approximation of the true absorption band features.

**Application to ROSIS images over Bourgneuf bay**

**Study area**

Bourgneuf Bay (46.47°N, 1.2°W) is located south of the Loire river and covers 340 km2. The intertidal zone extends over 100 km2. Sediment types range from sandy-mud in the south to mud in the north. Our study concerns only the northern part the foreshores that can be divided into rocky area and intertidal mudflats (Figure 6).
Figure 6: Schematic map of Bourgneuf bay. Location of the study area (rectangle).

Image Dataset
Reflective Optics System Imaging Spectrometer [ROSIS] images were acquired over Bourgneuf bay (SW France) in August 2002, in the framework of the HySens Campaign. ROSIS has 114 channels in the VNIR range [430-826 nm] with a spectral resolution of 4 nm and a pixel size of 2 m. Three stripes were necessary to cover our test area. They were georeferenced and a mosaic was built. Data were calibrated to radiance and converted to surface reflectance by DLR using the ATCOR4 atmospheric correction algorithm. Accuracy of the atmospheric correction was controlled using concurrent field reflectance spectra of reference targets.

Additional steps required for image application
In order to validate our automated process on a more general case, we applied the algorithm to ROSIS hyperspectral images. Some additional steps were required in order to take the greater diversity of spectral signatures into account. We first applied the MNF transform to eliminate most part of the noise from the data. Then, to determine a reasonable set of input parameters to start the procedure, Pixel Purity Index [PPI] and ND-Vizualizer were used to select spectral endmembers from ROSIS image. Each of these endmembers was run through the continuum removal and derivative analysis to estimate their spectral characteristics which are to be used as inputs for the MGM deconvolution. Finally, in order to find a mean continuum, an average spectrum is built from the endmembers and a reference set of parameters is built. The iterative MGM processing on the ROSIS image starts and the rest of the procedure is followed until the RMS reaches reasonable values.
Results
The spectra of the different constituents observable in Bourgneuf Bay are successfully modeled. The average RMS error is negligible 0.003 (stdev 0.01). Figure 7 shows examples of modeled spectra for different surface types.

Figure 7: Bourgneuf Bay ROSIS (RGB color associated with channels at 494, 586 and 674 nm). Green: microphytobenthos; light red: terrestrial vegetation; dark blue: macroalgae; white: sand, gray: mud and blue: water
Preliminary results show that pigment absorption for the biofilm (Chl-a and Chl-c for example) are successfully identified and modeled. It should also be possible to separate Fucus from macroalgae based on absorption positions and intensity. The interpretation of the various shapes of the continuum is less intuitive and requires additional work.

CONCLUSIONS AND PERSPECTIVES

These preliminary results show that an identification of sediments physical properties at the surface can be complicated. The MGM continuum removal seems more ‘realistic’ than the traditional continuum removal. The latter allows to make a comparison between reference and measured reflectance spectra and improve the identification. Using the MGM gives access to more quantitative information such as absorption depth, FWHM and area related to the concentration or fractional composition. Nevertheless it is more difficult to locate absorption features on raw reflectance spectra, particularly in the case of broad absorptions. Moreover, this technique is strongly sensitive to input parameters which require a priori knowledge. Several runs were necessary to adjust the initials parameters and reach a reasonable fit. In this study, we have shown that the spectroscopic derivative techniques coupled with MGM deconvolution allows to reduce a priori knowledge constraints. Consequently, the initials sets of the MGM parameters seem more robust. The spectroscopic derivative analysis gives us a better estimation of the number of absorptions, absorption wavelength position, local maxima and inflection points characterizing absorption bands. But, it is difficult to estimate the real number of overlapping and superimposed bands as well as their width. This algorithm is also time consuming when run on an image which shows great spectral variability.

Future work to improve algorithm performances includes: testing other band shape models such as Voigt or Lorentz functions; this could help solving the problem of “non realistic” overlapping bands for asymmetrical absorption bands and improve continuum adjustment.

Application to ROSIS data shows very satisfactory results that should allow quantitative analysis of the data and extraction of relevant physical parameters. Additional work is required now to precisely relate absorption and continuum characteristics to physical properties describing mudflat sediments. Spectral measurement performed in the laboratory under controlled environment should help better understand and constrain the relationships between spectral signatures and sediment parameters such as grain size or moisture content.

ACKNOWLEDGEMENTS

The authors would like to thank Mathieu Hirtzig (LPGN), Jean-Philippe Combe (SSI), Loïc Costeur (LPGN), for their helpful comments and fruitful discussions.
REFERENCES


