EPSC Abstracts
Vol. 10, EPSC2015-576, 2015
European Planetary Science Congress 2015

Online spectral fit tool (OSFT) for analyzing reflectance spectra

A. Penttilä (1), T. Kohout (1) and K. Muinonen (1,2)
(1) University of Helsinki, Finland, (2) Finnish Geospatial Research Institute (Antti.I.Penttila@helsinki.fi)

Abstract

We present an algorithm and its implementation for fitting continuum and absorption bands to UV/VIS/NIR reflectance spectra. The implementation is done completely in JavaScript and HTML, and will run in any modern web browser without requiring external libraries to be installed.

1. Introduction

The Online Spectral Fit Tool (OSFT) [1] was developed to our own needs in analyzing VIS-NIR spectral behavior of asteroids and meteorites. Our approach is similar to the Modified Gaussian Model (MGM) - software by Sunshine et al. [2, 3], but with certain differences in the model and in the implementation. We extended the applicability of our software by implementing it in JavaScript/HTML so that it can be used through a web browser. The OSFT runs in the client-side, and is self-consistent needing no plug-ins to be installed. The OSFT uses two public third-party JavaScript libraries [4, 5] included in the software.

We use no pre-determined spectral data, but fit the spectra using mathematical continuum model and analytical functions for absorption bands. The OSFT does not relate the fit directly to known spectral elements, but gives statistics describing the behavior of the continuum and the absorption bands. There are some semi-physical assumptions behind the MGM model. However, some of them are quite approximate when it comes to reflectance spectra of solid mineral samples. With the Central Limit Theorem, one can vaguely justify the Gaussian distribution as a model for an absorption band arising from several overlapping effects. On the other hand, also the Lorentzian band shape (i.e., Cauchy distribution) is used. Sunshine et al. [2] justify using the modified Gaussian model by crystal field theory and empirical tests. We feel that any of the abovementioned band shapes are not, strictly speaking, derived directly from physical principles applicable to mineral reflectance. Thus, we choose to use a band model that seems to fit best to our data.

2. OSFT Model Formulation

Our model for UV-VIS-NIR spectra is based on continuum with superimposed absorption bands. The model parameters are fitted using constrained non-linear optimization.

2.1. Absorption bands

We use the shape of the Gamma distribution for the absorption bands. The Gamma shape seems to fit the data better than the Gaussian, Lorentzian or modified Gaussian shape. The Gamma distribution is more peaked than the Gaussian, and is favored by our data. The distribution is defined for positive values, supporting its use with wavelengths. If the expected value μ of the Gamma distribution is large, i.e., $\mu/\sigma \gg 0$, the distribution is almost symmetric. We use an alternative parameterization of the distribution where μ and the standard deviation σ are the parameters of the distribution. The probability density function f is

$$f(x;\mu,\sigma) = \frac{1}{x\,\Gamma\left(\frac{\mu^2}{\sigma^2}\right)}\, e^{-\frac{\mu x}{\sigma^2}} \left(\frac{\sigma^2}{\mu x}\right)^{-\frac{\mu^2}{\sigma^2}}, \qquad (1)$$

where Γ is the Gamma function and $x > 0$. The actual absorption peak p in the model is the normalized Gamma distribution to have maximum value of 1[1], and multiplied by the peak strength parameter c:

$$p(x;\mu,\sigma,c) = \frac{c}{f(\mu;\mu,\sigma)}\, f(x;\mu,\sigma). \qquad (2)$$

The area A of the band can be derived analytically:

$$\begin{aligned} A(\mu,\sigma,c) &= \int_0^\infty p(x;\mu,\sigma,c)\,dx \\ &= c\, e^{\frac{\mu^2}{\sigma^2}}\, \mu \left(\frac{\mu}{\sigma}\right)^{-\frac{2\mu^2}{\sigma^2}} \Gamma\left(\frac{\mu^2}{\sigma^2}\right), \end{aligned} \qquad (3)$$

[1] Approximately. Because the distribution is not strictly symmetric, $f(\mu;\mu,\sigma)$ can be slightly smaller than the maximum value of the distribution.

while the full-width-at-half-max (FWHM) w of the peak needs to be solved numerically by finding x's.

2.2. Continuum

We wanted the continuum model to be more flexible than the first-order polynomials. For this reason, we chose cubic splines as the basis of the continuum. The continuum is modeled as the linear combination of five B-spline basis functions of third degree with knots in the range from the smallest wavelength λ_{min} to the largest wavelength λ_{max} in the data. The coefficients b_i of the linear combination are to be fitted.

2.3. Fit

The final fit $g(\lambda)$ to the data is:

$$g(\lambda) = \sum_{i=0}^{4} b_i s_i(\lambda) - \sum_{i=1}^{n} p(\lambda; \mu_i, \sigma_i, c_i), \quad (4)$$

where s_i are the ith B-splines of degree 3, λ is the wavelength, n is the number of absorption peaks, and μ_i, σ_i, c_i are the mean, standard deviation, and depth for peak i.

2.4. Optimization

The function $g(\lambda)$ is fitted to the data using constrained non-linear optimization. The sum-of-squared-errors (SSE) between the data and the model is minimized. Constraints need to be set for the band parameters. All standard deviations and depth parameters $\sigma_i, c_i \geq 0$. The mean parameters need to be ordered, $\lambda_{min} \leq \mu_1 \leq \cdots \leq \mu_n \leq \lambda_{max}$. Actually, in addition to minimizing SSE, we also introduce two other penalizations with small weights. These penalizations are related to the continuum being smooth and the bands to extend only to the area of absorption and not to continuum.

3. Usage

The beta version of the OSFT is publicly available [1]. The user can upload a text-file of wavelength-reflectance pairs. The data is shown graphically, and the user can select the number of absorption bands. Initial parameter guesses need to be given for band centers, deviations, and depths. With initial values for bands, the continuum is fitted using the linear least-squares fit, and the initial fit is displayed with the data. Once the initial fit is in the right neighborhood, the algorithm will do the actual non-linear fit for the band and continuum parameters together. When completed, the software can compute fit statistics, i.e., band locations, deviations, depths, FWHM-values, and band areas, together with the slope and albedo parameters. At this point, the data can also be normalized at a given wavelength. We prefer to normalize only after the fitting, since we rather normalize with the value from the fitted continuum than with a single measured reflectance value. A single measured value can contain errors which will affect the slope and band depths, while the fitted continuum is 'an average' and thus have smaller errors. Fig. 1 shows two examples of the fit produced by the OSFT.

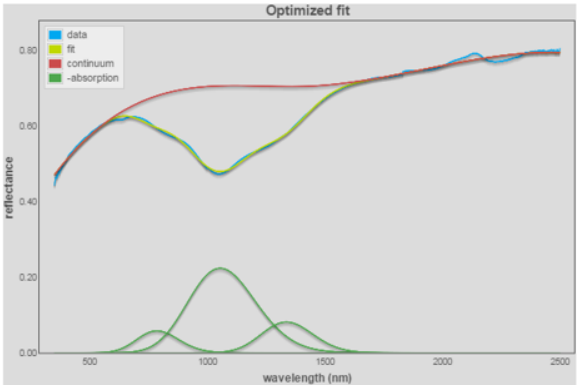

Figure 1: OSFT fit with three overlapping absorption bands for olivine including nanophase Fe^0.

Acknowledgements

AP and KM acknowledge the ERC Advanced Grant No. 320773, Scattering and Absorption of Electromagnetic Waves in Particulate Media (SAEMPL).

References

[1] A. Penttilä and T. Kohout (2015). OSFT tool, http://www.helsinki.fi/project/psr/OSFT/.

[2] J.M. Sunshine, C.M. Pieters, and S.F. Pratt (1990). JGR 95, 6955–6966.

[3] J.M. Sunshine, C.M. Pieters, S.F. Pratt, and K.S. McNaron-Brown (1999). Abstract, LPSC 30.

[4] A. Gustafsson (2012). Cobyla2 port to JavaScript, http://code.google.com/p/jscobyla/.

[5] Flotr2 Canvas graphing library (2014), http://www.humblesoftware.com/flotr2/.