

ANALYSIS OF MATHEMATICAL MODELS FOR REFLECTANCE ABSORPTION SPECTRA. I. Antonenko¹ and E. A. Cloutis², ¹Planetary Institute of Toronto, 197 Fairview Ave. Toronto, ON M6P 3A6, CANADA, ²Department of Geography, University of Winnipeg, 515 Portage Ave., Winnipeg, Manitoba R3B 2E9, CANADA.

Introduction: Reflectance spectroscopy is a useful planetary exploration technique, particularly for quantitative geological mapping of planetary surfaces [e.g. 1]. However, its full potential has not yet been realized because of two major limitations. 1) Hyperspectral imagers can acquire more data than accommodated by available data transmission techniques, even with existing compression algorithms. As a result, trade-offs in terms of spectral resolution vs spatial resolution vs temporal coverage often occur. 2) Current spectrum analysis algorithms are labour-intensive and require user intervention to perform tasks such as baseline fitting and optimization of number of absorption bands to be modeled [e.g., 2, 3]. For this reason, many spectral data sets have not been fully analysed, even years after their acquisition (e.g., Phobos 2 ISM data).

Both these limitations could be addressed by novel algorithms that simultaneously provide automated analysis of hyperspectral data (yielding information relating to end member abundances and compositions) as well as an order of magnitude reduction in data volume. Such algorithms could be implemented on board spacecraft, allowing for the development of intelligent hyperspectral sensors, which are programmed to seek out and focus on geologically (and biologically) interesting regions on planetary surfaces such as Mars.

As a first step to such an automated process, we are investigating the mathematics of absorption features. At present there is no consensus concerning which functions best represent absorption bands. Candidates include Gaussian, modified Gaussian, and Lorentzian functions [e.g., 3, 4]. In this study, we attempt to determine which functions are appropriate for which type of feature (e.g. crystal field transitions vs intervalence charge transfers) by using an empirical approach.

Method: In this first phase of our investigations, we confined our study to pyroxene spectra, addressing crystal field absorptions at 1 & 2 μm , and charge transfer absorption at ~ 800 nm. A selection of pyroxene absorption spectra, acquired at NASA's RELAB facility at Brown University [5], were chosen because of their well-defined spectral properties, compositions and structures. These spectra were analyzed using a commercial software package called PeakFit (version 4.0).

Each spectrum was converted to a modeling space of wavelength (cm^{-1}) vs. inverse natural log reflectance. This allows spectra to be modeled as linear combinations of mathematical functions, lets baselines be modeled by straight lines [3], and allows absorption fea-

tures to be modeled as peaks. A digital filter was applied to give the data uniformly spaced x-values, which are necessary for many of the PeakFit algorithms. The data was sectioned, so that only the peaks of interest (0.8, 1, & 2 μm absorption features) were included for fitting. Sectioned data included all points from local minimum to local minimum at either end of the area of interest. The data set was then smoothed, using the Loess procedure, a least-squares weighted regression smoothing algorithm.

Sectioned data was fit using PeakFit's automated procedure based on residuals. The program provided initial estimates of peak location and number. Number of initial peaks was adjusted manually, based on visual analysis, so the correct number of peaks was present. The PeakFit algorithm was then iterated until it converged on a solution. Peak widths were allowed to vary independently during the fitting. A linear baseline, determined initially by the 2nd zero derivative, was also fit during the peak modeling. A solution was found when improvement in root mean square error (RMS) was negligible (on the order of 10^{-5} or less) over several successive iterations.

Results: Several mathematical functions were fit to our spectra sections: Gaussian, modified Gaussian (MGM), and Lorentzian functions, as well as the Voigt function (a convolution of the Gaussian and Lorentzian functions). Each function was fit to each absorption feature in a spectral section. The results were assessed visually and the RMS noted (Figure 1). Good fits were visually indistinguishable, so best function for each peak was determined from lowest RMS.

All spectral sections had >1 peak present. Though each peak was assigned a function independently, lowest RMS was considered for all peaks together (Figure 1). For several spectra, we also fitted each absorption feature independently, creating sections with only single peaks. Although these results are not as robust (because they produced different baseline fits for each feature) they generally return the same peak features and best mathematical function. This is particularly true when peaks are well separated and can be modelled independently. In many cases, peaks are close and interact at the wings. In these cases, attempts to fit peaks individually do not produce the same results as when all the peaks in a spectrum are fit together.

Contrary to expectations, a number of different mathematical functions were found to model most feature types. We plotted peak centre vs peak width, indi-

cating which mathematical function best modeled the peak (Figure 2). Some generalizations can be made. In all cases, the 800 nm feature was best modeled by the MGM function. The $2\mu\text{m}$ band was modeled by both Gaussian and MGM functions. The $1\mu\text{m}$ band was modeled by the Gaussian, MGM and Lorentzian functions. There appears to be some correlation between function and band width. In the $2\mu\text{m}$ region, MGM-modeled peaks have wider widths than Gaussian peaks. In the $1\mu\text{m}$ region, MGM peaks tend to be wider than Gaussian peaks, and Lorentz peaks are wider still. Reasons for this are unclear and requires additional study.

Only two peaks were found to be modeled by the Voight function. Both are from the same spectrum (PYX150), a clinopyroxene with 3 peaks in the $1\mu\text{m}$ region. The centre peak is modeled by a Gaussian, and the two flanking peaks by Voight functions. Whether this is significant or is simply a poor fit of 3 closely spaced and overlapping features is not clear at present.

Findings relevant to automation: In our investigations, we noticed that the PeakFit solution is insensitive to initial conditions. We tested this theory and found that variations in starting position of baseline and peaks do not affect the final peak characteristics.

PeakFit offers 3 algorithms to find starting condition of peaks. Again, the solution is insensitive to algorithm used, because initial conditions are irrelevant.

However, the number of initial peaks can be very important. Too many peaks at the start allow for solutions with very low RMS, but no physical significance. Generally, excess peaks will have their amplitude decreased until they no longer significantly. An amplitude rejection threshold can be applied to eliminate such peaks. Other peaks migrate out of the region of interest and can be eliminated by a range threshold. Occasionally, excess peaks find a stable configurations that gives a very good fit, both visually and using RMS, but have no physical validity (e.g. 2 equally sized peaks on either side of feature centre). These cases require intervention, complicating the automation process.

During fitting, we found that final peak centres were somewhat insensitive to the mathematical function used. Variations in peak centres between different peak models used are confined to 10nm or less.

Conclusions: A preliminary analysis of mathematical functions, as applied to reflectance absorption features, suggest several trends. For pyroxenes, the 800nm region is modeled best by the modified Gaussian model (MGM). The $2\mu\text{m}$ region is modeled by Gaussian and MGM functions. Gaussians appears to predominate at narrower peak widths. The $1\mu\text{m}$ region has a more complex relationship between function type and peak width. We plan to extend our study to other minerals, to see if these trends hold up.

During this study, we identified several factors that could assist in automation of spectral analysis. The PeakFit algorithm is relatively insensitive to the starting conditions of peak and baseline location. Human involvement in baseline fitting can potentially be eliminated. Unfortunately, optimization of the number of peaks is still an issue. Also, location of the fitted peak centres is relatively stable, regardless of functions used. Thus choice of function may not be as important as originally thought, at least in regards to peak centre.

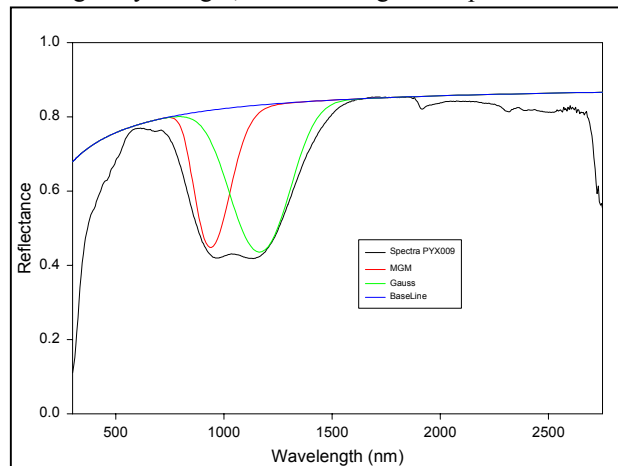


Figure 1: Spectrum of PYX009 and best fit mathematical functions. RMS for this fit was 0.0017707.

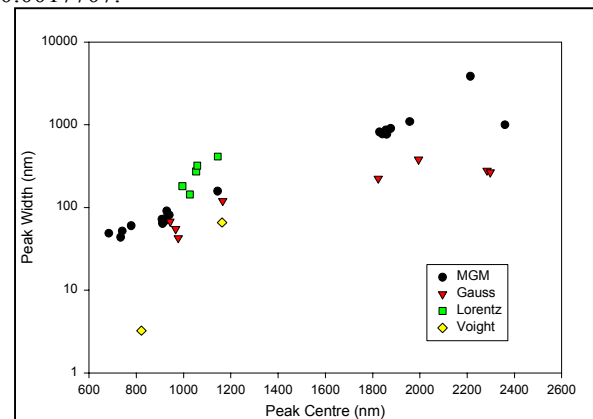


Figure 2: Graph of peak centre vs peak width, showing trends between peak and best fit function.

References: [1] Bell and Crisp, 1993, *Icarus* **104**, 2-19; Bell et al., 1997, *Bull. of the Am. Astron. Soc.* **29**, 963; Erard et al., 2000, *LPSC* **31**, abstract #1325; Kirkland et al., 2000, *JGR* **105**, 22,507-22,515. [2] Johnson et al., 1992, *JGR* **97**, 2649-2657; Murchie et al., 1993, *Icarus* **105**, 454-468. [3] Sunshine et al., 1990, *JGR* **95**, 6955-6966. [4] Burns, 1970, *Min. App. of Crystal Field Theory*; 224 pp., Cambridge U. Press, New York; Clark and Roush, 1984, *JGR* **89**, 6329-6340. [5] Pieters 1983, *JGR* **88**, 9534-9544.

This study was supported by a research contract from ESA-ESTEC through the Aurora program