Computers & Geosciences 100 (2017) 103–114

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Review article

Fitting the curve in Excel®: Systematic curve fitting of laboratory and remotely sensed planetary spectra

Michael A. McCraig[a,*], Gordon R. Osinski[a,b], Edward A. Cloutis[a,c], Roberta L. Flemming[a], Matthew R.M. Izawa[c,d], Vishnu Reddy[e,1], Sherry K. Fieber-Beyer[f,1], Loredana Pompilio[g], Freek van der Meer[h], Jeffrey A. Berger[a], Michael S. Bramble[a,i], Daniel M. Applin[c]

[a] Department of Earth Sciences/Centre for Planetary Science and Exploration, University of Western Ontario, 1151 Richmond St N, London, Ontario, Canada N6A 5B7
[b] Department of Physics and Astronomy, University of Western Ontario, 1151 Richmond St N, London, Ontario, Canada N6A 5B7
[c] Department of Geography/Hyperspectral Optical Sensing for Extraterrestrial Reconnaissance Laboratory, University of Winnipeg, 515 Portage Ave., Winnipeg, Manitoba, Canada R3B 2E9
[d] Institute for Planetary Materials, Okayama University, 827 Yamada, Misasa, Tottori 682-0193, Japan
[e] Lunar & Planetary Laboratory, The University of Arizona,1629 E University Blvd., Tucson, AZ 85721-0092, USA
[f] Department of Space Studies, University of North Dakota, University Stop 9008, Grand Forks, ND 58202, USA
[g] Department of Psychology, Health, and Earth Sciences (DiSPUTer), D'Annunzio University, 66100 Chieti, Italy
[h] University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), Hengelosestraat 99, 7514 AE, Enschede, The Netherlands
[i] Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, RI 02912, USA

ARTICLE INFO

Keywords:
Data reduction techniques
Curve fitting
Spectroscopy
Spectral analysis
Curve deconvolution
Modelling
Asteroids
Mineralogy
Mars

ABSTRACT

Spectroscopy in planetary science often provides the only information regarding the compositional and mineralogical make up of planetary surfaces. The methods employed when curve fitting and modelling spectra can be confusing and difficult to visualize and comprehend. Researchers who are new to working with spectra may find inadequate help or documentation in the scientific literature or in the software packages available for curve fitting. This problem also extends to the parameterization of spectra and the dissemination of derived metrics. Often, when derived metrics are reported, such as band centres, the discussion of exactly how the metrics were derived, or if there was any systematic curve fitting performed, is not included. Herein we provide both recommendations and methods for curve fitting and explanations of the terms and methods used. Techniques to curve fit spectral data of various types are demonstrated using simple-to-understand mathematics and equations written to be used in Microsoft Excel® software, free of macros, in a cut-and-paste fashion that allows one to curve fit spectra in a reasonably user-friendly manner. The procedures use empirical curve fitting, include visualizations, and ameliorates many of the unknowns one may encounter when using black-box commercial software. The provided framework is a comprehensive record of the curve fitting parameters used, the derived metrics, and is intended to be an example of a format for dissemination when curve fitting data.

1. Introduction

Spectroscopy is the principal – and sometimes the only – technique that can provide information regarding the compositional and mineralogical make up of planetary bodies. Notable examples include the linkage between the ordinary chondrites and the S-type asteroids cemented by the Itokawa – Hayabusa encounter and verified by sample return (Binzel et al., 2001; Abe et al., 2006; Abell et al., 2007; Nakamura et al., 2011); the link between the Howardite-Eucrite-Diogenite (HED) meteorites and asteroid (4)Vesta (McCord et al., 1970; Gaffey et al., 1982; Gaffey, 1983, 1997), confirmation of which was provided by the Vesta-Dawn encounter (McSween et al., 2013); and the soon to fly OSIRIS-REx sample return mission which is expected to confirm the link between the carbonaceous meteorites and the missions target, B-type asteroid Bennu [1999 RQ_{36}] (Clark et al., 2011; Hergenrother et al., 2013). Given the robustness of the connection between observational asteroid spectra and their analogous meteorites confirmed thus far, it is also rather likely that the proposed

* Corresponding author.
E-mail address: michaelmccraig@hotmail.com (M.A. McCraig).
[1] Visiting Astronomer at the IRTF, which is operated by the University of Hawai'i Mauna Kea under Cooperative Agreement NNX-08AE38A with the National Aeronautics and Space Administration, Science Mission Directorate, Planetary Astronomy Program.

http://dx.doi.org/10.1016/j.cageo.2016.11.018
Received 5 January 2016; Received in revised form 14 September 2016; Accepted 30 November 2016
Available online 02 December 2016
0098-3004/

linkages such as those suggested for the Chelyabinsk meteorite and Baptistina asteroid family by Reddy et al. (2014), the mesosiderites and the Maria asteroid family suggested by Fieber-Beyer et al. (2011), and the relationship between (6) Hebe and the H-type ordinary chondrites suggested by Gaffey and Gilbert (1998), are similarly as robust. We use spectroscopy to remotely probe a wide variety of solar system bodies that we cannot visit, primarily using reflectance spectroscopy in the ultraviolet through near-infrared wavelengths and emission spectroscopy in the mid through far-infrared. To discern mineralogy using spectroscopy we attempt to match the unknown with known spectra from spectral libraries or diagnostic spectral features and to do this we must curve fit the collected spectra.

Despite all of the obvious successes, and a raft of literature on the subject, there is what can be best described as a communication problem within the planetary science spectroscopic community. The primary metrics one needs to derive and communicate are the positions of the maxima/minima/centres for absorption bands and secondarily the band areas/Band-Area-Ratios (BAR) and band depths which are routinely included in manuscripts, but often presented with no reference to the measurement methods or the constraints on the curve fits used to derive them. If the measurement methods are mentioned, for those new to the field, it can be confusing what is meant when an author refers to the application of straight- or curved-line continua, a method of apparent continuum removal or the centring of absorption bands. Often causing further confusion is the use of the word modelling to describe a number of analytical methods which meet a broad definition of the word "modelling," though most often are not modelling of the physical processes responsible for the absorptions. While these methods are technically modelling, the fitting of curves, e.g., curve deconvolution using Gaussians fit to raw or apparent continuum removed reflectance spectra, cannot model the physical processes which are responsible for the creation of spectral absorptions in energy space and should aptly be referred to as Gaussian fitting rather than modelling.

When attempting to describe an object with spectral data, researchers should define the curve fitting algorithms and the fitting parameters used, and the constraints and uncertainties of the fits. We need to ensure that useful data is being disseminated and derived metrics are mathematically defined for repeatability and validation.

Metrics should be derived from spectra via a process of measurement of absorption band positions, i.e., minima, maxima and centres, as well as absorptions depths and when applicable, areas, skew, and so on, using a standardized methodology for curve fitting that is entirely transparent and fully repeatable. In this work we have assembled as many of the spectral analysis and deconvolution functions we could envisage that would be useful, and those that could be useful for furthering understanding and interpretation of, and extraction of information from, spectral data, as a set of what should be universally applicable functions with suggestions for implementation to curve fit spectra.

The included supplementary materials in Appendices A and B are a guide to robust and reliable methods for empirical curve fitting that, first, demonstrate the individual mathematical functions used for those who may be new to spectroscopy and may be having difficulty visualizing and performing curve fitting procedures such as straight-line apparent continuum removal, minima/centre derivation and so on. And secondly, we believe the examples in the Appendices are good illustrations of the necessity for the dissemination of more completely defined parameterized data to the community.

The concept of curve fitting spectra (also often referred to as spectral deconvolution) is not new, and continues to evolve (e.g., Doetsch, 1928; Lonn, 1932; Kaper, 1966; Clark and Roush, 1984; Sunshine et al., 1988; van der Meer, 2004; Clenet et al., 2011; Parente et al., 2011). Curve fitting is a necessity as all measured spectroscopic data are a series of data points employing x and y coordinates for localization in wavelength or energy space. One has to remember that, while the actual waveform is analog, and spectra are often visually represented by a solid line, collected spectra are not lines or curves, rather, they are a collection of discrete points which subsample the analog waveform. As such, we fit curves representative of the actual waveform through the measured discrete points, along which any number of corresponding x and y points can be interpolated to either calculate a wavelength or wavenumber value for the absorption band minimum/maximum/centre, or discern which measured point in the existing data is closest to what would be the perfectly resolved interpolated minimum/maximum/centre.

The fitting of mathematical functions to a spectrum is a method that allows the derivation of a set of metrics that are used to describe properties of a spectrum and which can be used to decipher the mineralogy or mineral assemblage responsible for the features of a particular spectrum. This system needs to be transparent, repeatable, and ideally easy to implement, such that the derivation of the reported metrics is eminently clear. Spectra obtained from planetary surfaces do not have the set of constrained variables that prepared laboratory samples have. In the laboratory, grain size, packing (and to an extent porosity), temperature, phase angle (i.e., viewing geometry) and signal-to-noise-ratio can be controlled, and mineral mixtures can be precisely constrained or understood. In addition, given the ever present complications caused by processes operating on planetary surfaces (such as space weathering), several of the metrics that might allow one to curve match one spectrum with another such as slope, band depth and albedo can differ significantly between planetary surfaces and laboratory spectra for even small differences in mineral or mixture composition (e.g., Gaffey, 2010). The most important metrics are arguably the band minima/maxima and the centre as they are the most sensitive to mineralogy and the least sensitive to the spectral effects of grain size, packing, temperature and the poorly constrained effects of space weathering one must deal with for atmosphere-less bodies (Gaffey, 2008).

2. Curve fitting methodologies

This work is focused on the curve fitting of ultraviolet through near-infrared reflectance spectra of minerals, asteroids and planetary surfaces for the purpose of interpreting remote-sensing spectra of planetary bodies – though many of the methodological points are equally applicable to any other spectrum or curve one might want to fit or model.

In order to successfully curve-fit spectra of solid minerals, a certain amount of *a priori* knowledge is required of spectroscopy, mineralogy, and the basics of curve fitting mathematics. In order to understand how to effectively curve fit spectra it is important to understand the drivers for spectroscopic absorptions and be familiar with the lexicon. One also needs to understand the limitations of the mathematics involved.

Researchers use varying combinations of techniques to derive qualitative or quantitative information from spectra, e.g., curve matching, curve deconvolution, and empirical curve fitting. Each technique has strengths and weaknesses, and an understanding of the application of one or more techniques requires precise knowledge of the spectroscopic lexicon and the metrics the various techniques provide.

To understand the majority of the spectroscopic absorptions occurring in ultraviolet through infrared reflectance spectra of planetary and asteroidal solid surfaces, one operates from the hypothesis that all absorptions occur at a specific central maximum or minimum in energy space, the probability of absorption of a photon, which translates into the shape of an absorption feature (or band) will follow a normal symmetric distribution in energy space about its centre, and the 'wings' on either side of the absorption extend to infinity (assumed band symmetry is a minor oversimplification for the purposes of this discussion). If an absorption band is not symmetric, the asymmetry must be due to another contributing absorption, either from the same material or another constituent in the assemblage, a continuum that