

## Review article

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



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

<sup>a</sup> Department of Earth Sciences/Centre for Planetary Science and Exploration, University of Western Ontario, 1151 Richmond St N, London, Ontario, Canada N6A 5B7

<sup>b</sup> Department of Physics and Astronomy, University of Western Ontario, 1151 Richmond St N, London, Ontario, Canada N6A 5B7

<sup>c</sup> Department of Geography/Hyperspectral Optical Sensing for Extraterrestrial Reconnaissance Laboratory, University of Winnipeg, 515 Portage Ave., Winnipeg, Manitoba, Canada R3B 2E9

<sup>d</sup> Institute for Planetary Materials, Okayama University, 827 Yamada, Misasa, Tottori 682-0193, Japan

<sup>e</sup> Lunar & Planetary Laboratory, The University of Arizona, 1629 E University Blvd., Tucson, AZ 85721-0092, USA

<sup>f</sup> Department of Space Studies, University of North Dakota, University Stop 9008, Grand Forks, ND 58202, USA

<sup>g</sup> Department of Psychology, Health, and Earth Sciences (DiSPUTer), D'Annunzio University, 66100 Chieti, Italy

<sup>h</sup> University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), Hengelosestraat 99, 7514 AE, Enschede, The Netherlands

<sup>i</sup> 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<sub>36</sub>] (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](mailto:michaelmccraig@hotmail.com) (M.A. McCraig).

<sup>1</sup> 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/© 2016 Elsevier Ltd. All rights reserved.

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

superimposes an asymmetry on a symmetric absorption, or the absorption band may be saturated, which can make an absorption appear to be flattening across its minimum or otherwise affect the apparent symmetry of a band.

Curve matching is a catch-all term for a number of practices where one is attempting to compare one spectrum with another, mostly visually. The most common application uses a technique where one employs a library of sample spectra, taken from laboratory data of usually pristine, monomineralic, prepared (i.e., powdered, sieved, washed, etc.) samples which are combined mathematically in a linear fashion (often called spectral deconvolution, linear deconvolution and/or quantitative deconvolution, not to be confused with curve deconvolution), e.g.,  $a^{0.5} + b^{0.25} + c^{0.25} = x$ , to produce a facsimile spectrum which closely matches the entire spectrum of an unknown sample. The facsimile spectrum is overlain atop the observed spectrum of an unknown sample or surface, and modified by either varying amounts of individual contributors, e.g.,  $a^{0.25} + b^{0.5} + c^{0.25} = x$ , and or adding or subtracting more contributors, e.g.,  $a^{0.25} + b^{0.25} + c^{0.15} + d^{0.1} + e^{0.05} + f^{0.05} + g^{0.05} + h^{0.05} + i^{0.03} + j^{0.02} = x$ , until an acceptable match with the observed spectrum is produced which appears analogous to the person performing the curve matching or until some threshold residual for the facsimile versus observed spectrum is met (e.g., [McSween et al., 2003](#), and references therein). As a first pass, the technique can be illustrative if one has very little context for the spectrum under investigation but it requires a spectral library and algorithms (e.g., [Ramsey and Christensen, 1998](#)) to produce and fit the deconvolved spectrum/facsimile to some residual or human input to suggest or rule out contributors, and it does not capture the fact that spectra of mineral assemblages rarely, if ever, mix linearly ([Singer, 1981](#); [Clark, 1999](#); [Kraft et al., 2003](#); [Berger et al., 2015](#)). As a technique it may also ignore other factors that can contribute to altering a particular mineral, or an assemblage of mineral spectra, such as, grain size, porosity, texture or roughness, grain packing, temperature, phase angle, atmosphere, adsorbed species, spatial resolution, spectral resolution, distance, the probable effects of space weathering, and so on. This is problematic as libraries of spectra that explore alteration of spectra by factors like phase angle and grain size do not exist in any quantity. Curve matching can also refer to searching databases of spectra for samples with absorptions/curves/slopes that are similar to a sample's or simply plotting multiple spectra on the same chart to compare one spectrum with another and noting plausible similarities, or suggesting similarities with vertical lines to draw the eye. No matter the methodology, as a set of techniques to constrain actual mineralogy, curve matching often fails to provide robust quantitative information (e.g., [Gaffey, 2008, 2010](#)). This simple technique can be useful for initially constraining possible spectral contributions, which can then be optimized for subsequent more rigorous approaches.

Curve deconvolution/Gaussian fit optimization is a quantitative approach based on the mathematics of spectral absorptions using Gaussian fitting for solid material spectra. It correctly operates under the supposition that the majority of absorptions are symmetric in energy space, and any band asymmetries are contributions from an adjacent absorption(s). Importantly though, this technique can fail to converge or provide realistic results if some initial constraints are not placed on the fitting algorithms, i.e., approximate starting centre wavelengths, and the number of absorptions to use for a particular fit. It may also fail if more than a single monomineralic sample spectrum is fed into the deconvolution algorithm and initial parameterization does not take this into account, though researchers are presently working on methodologies to remove the need for initial fixed parameter sets (e.g., [Makarewicz et al., 2009](#), [Clenet et al., 2011](#); [Buz and Ehlmann, 2014](#)). The most widely used and freely available deconvolution software package used in planetary science is the Modified Gaussian Model (MGM) suite of scripts for MATLAB™ produced by [Sunshine et al. \(1988, 1989, 1990, 2004\)](#) and [Sunshine](#)

[and Pieters \(1993\)](#). If applied correctly, curve deconvolution (more aptly called, Gaussian fit optimization) provides a quantitative approach, but 'correct' application requires the spectrum under investigation to be reasonably well known to constrain the initial starting parameters, e.g., one must constrain an MGM analyses by directing MGM to optimize the fit of two Gaussians for the first and one Gaussian for the second of the two large near-infrared absorptions present in pyroxene. In [Sunshine et al. \(1989\)](#) and [Sunshine and Pieters \(1998\)](#), the authors address the problems that arise with a set of unconstrained variables, noting that the number of absorptions and wavelengths for the band centres of those absorptions must be in the initial parameter set, and should they not be provided because the sample mineralogy is unknown, there are simply too many unconstrained variables and any curve deconvolution/Gaussian fit optimization techniques will not produce nonunique solutions. Curve deconvolution as a technique has real merit and may be the best technique for reliably identifying a shoulder on an absorption (i.e., a smaller absorption band visible on the wing of a larger absorption band).

It should also be noted that [Gallie et al. \(2008\)](#) found that for the purposes of fitting, Gaussians in wavenumber, i.e., energy space, and those in wavelength space are numerically equivalent. So, if one is attempting to fit a series of Gaussians, it matters little whether one is working in energy space or wavelength space, though there will always be an issue of repeatability as automated fit optimization as performed by MGM, and the methods demonstrated in the [Appendices](#) are exceedingly unlikely to produce the same grouping of Gaussians there, producing unrepeatability (see [Appendix A](#)). One also has to be aware that while MGM has the word "Model" in its name, it is not in normal practice modelling absorptions, as the standard procedure is to remove straight line apparent continua a reflectance spectrum, converted to energy space, which has already had its imposed continuum removed.

An alternative to curve deconvolution is empirical curve fitting, the technique used and advocated for herein. In this method, spectra are curve fit systematically using polynomials fit about the minima or centres of absorptions to derive said metrics, and then the fit minima/centres can be compared with similar curve fits performed using laboratory spectra of calibrated standards and calibrated mixtures. The methodology uses a simple set of rules for curve fitting where one does not need to worry about unconstrained/over-constrained/under-constrained variables, fit residuals or the intricacies of the application of Gaussian fit optimization/curve deconvolution. In its simplest application, such as when fitting hydrated phyllosilicates hydration features for comparison with Martian remote sensing spectra, after converting Martian intensity-over-flux measures (I/F) to reflectance spectra as per [Bakker et al. \(2014\)](#), the user would do no more than fit absorption minima/maxima and provide those numbers with their manuscript as per the examples in [Appendix A](#). Ideally, full fits, including spectra would be included as [Supplementary material](#) with any published manuscript so no questions arise later about what was done when the spectra were curve fit.

For decades, a number of researchers and research groups have advocated for this methodology (e.g., [Kaper, 1966](#); [Gaffey, 1976, 2010](#); [Clark, 1980, 1981, 1999](#); [Farr et al., 1980](#); [McCord et al., 1981](#); [Singer, 1981](#); [Clark and Roush, 1984](#); [Cloutis et al., 1986](#); [Cloutis and Gaffey, 1991](#); [Gaffey et al., 1993, 2002](#); [Gaffey and Gilbert, 1998](#); [Gaffey and McCord, 1977, 1978](#); [Clark et al., 2003](#); [van der Meer, 2004](#); [Storm et al., 2007](#); [Burbine et al., 2009](#); [van Ruitenbeek et al., 2014](#)), largely because it is easy to apply and reproduce when the fitting procedure is well defined. Researchers often use proprietary routines that perform very similar or identical functions to those we have provided herein, and/or the SPECTrum Processing Routine software (SPECPR). For the purposes of empirically fitting spectra, we are interested in applying and performing operations that are identical to the straight-line apparent continuum removal of SPECPR and Tetracorder, and the polynomial fitting functions of SPECPR ([Clark, 1980, 1993](#); [Clark et al., 2003](#); [Livo and Clark, 2014](#)); software available from the USGS).

SPECPR and Tetracorder are freely available as are instruction manuals, but the software can be extremely cumbersome to use.

### 3. Empirical curve fitting

Empirical curve fitting relies on a series of calculations used to derive a set of metrics from a spectrum that can be used for characterization and comparison (e.g., Singer, 1981; Clark and Roush, 1984; Cloutis and Gaffey, 1986). As noted, when deriving these metrics, *a priori* knowledge of the physical processes that produce spectroscopic absorptions is ideal, but not a requisite. A central guiding principle is that all absorptions are distributed about a specific central point in energy space and that the absorption feature, or band, will follow a Gaussian distribution of energies around that centre (in solids), and the 'wings' on either side of the absorption go 'to infinity' – that is, they extend with ever decreasing intensity to higher and lower energy levels than the band centre. If a band does not conform, i.e., it is not symmetric, the asymmetry must be due to another contributing absorption, either from the same material, another constituent in the assemblage under investigation, or a continuum that superimposes asymmetry on a symmetric absorption, or band saturation is occurring (Clark and Roush, 1984; Lucey and Clark, 1985.) The symmetry of an absorption as a central tenet applies to spectra in energy space but not to spectra in wavelength where the majority of the metrics we are attempting to derive originate, so when empirically curve fitting we can somewhat ignore the drivers for the absorptions we are fitting.

In order to derive some meaning from a single band, or set of bands, some knowledge of the possible structures contributing to the absorption is required but we are not attempting to account for everything producing the spectrum when deriving metrics empirically. For instance, if one were trying to derive a set of metrics from the laboratory spectrum of the pyroxene sample shown in Figs. 1 through 4, knowing that pyroxene has absorptions due to different electronic interactions with  $Fe^{2+}$  in the M1 and M2 octahedral sites, producing the absorption feature referred to as Band I, and is a combination of two absorptions at  $\sim 1$  and  $1.15 \mu m$  while Band II at  $\sim 2 \mu m$  is a singular feature, is advantageous, but not a requisite (Cloutis et al., 1986; Burns, 1993; Klima et al., 2007, 2011). Given the limitations of the curve fitting techniques described herein (or any other techniques, see Appendix A and B), how one would logically go about deriving absorption band minima, centres, depths and areas would be driven by the directives for application and limitations of the method/technique, *a priori* knowledge of the causes and number of absorption features and how a set of derived metrics would compare with those gathered from pre-existing research. For example, while the second contributing absorption of Band I in the pyroxene example spectrum is obvious, its exact position and contribution to the shape and area of Band I is not

as important as understanding that it should be included in centring, depth and area determination. This is because the directive in empirical curve fitting is to include its contribution by fitting straight-line apparent continua to the point of band extinction (the point of maximum reflectance between two adjacent absorption band minima), therefore it is included in the metrics one is producing now just as it has been in metrics produced in the past by other researchers employing the same techniques.

Despite methodological differences, from all spectra one can derive the main metrics of an absorption band or feature, i.e., the minimum, band centre, band depth and band area; and other less critical metrics such as measures of asymmetry, i.e., skew, kurtosis; measures of width, i.e., full width at half maximum (FWHM), and area, that should have universally applicable definitions, but often do not. Below, we provide definitions of the most important terms.

#### 3.1. Band minima

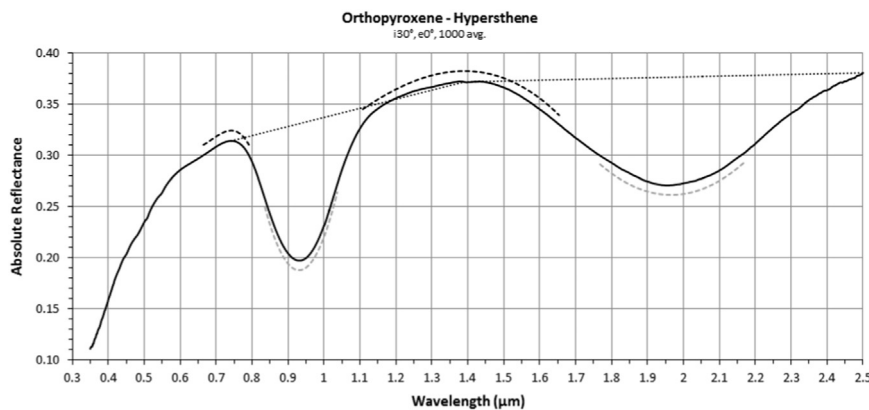
Denotes the wavelength or energy of the position of lowest reflectance over a specified wavelength interval, derived from a spectrum that has not had its imposed continuum removed. There should always be a definition of the procedure for measuring the band minimum which includes a description of any modifications made to the spectra before the minima are measured, e.g., breaks in the spectra at the detector junctions that may have been removed by application of a particular function, the reflectance spectra was fit as delivered by the spectrometer, the spectrum was offset or normalized, etcetera.

#### 3.2. Band centre

Denotes the wavelength position of lowest reflectance over a specified interval after some kind of continuum removal for a band has been performed; e.g., apparent continuum removal by division, i.e.,  $y$  divided by the straight or curved line apparent continuum value for the same  $x$  value.

#### 3.3. Band depth

Is defined in a number of ways, but it is always the depth of an absorption at the derived wavelength minimum or centre value relative to a defined point above that value. Most often, band depth is the length of a vertical line expressed in percent reflectance, which extends downward tangentially from 100% reflectance or unity, to the derived centre or minima (e.g., Clark and Roush, 1984; Clark, 1999). Band depth is normally expressed in percent reflectance, even in cases where bands have been contrast stretched by imposed continuum removal and the centring process of further apparent continuum removal.



**Fig. 1.** The NIR spectrum of the pyroxene, solid black line, with black offset (+0.01) dashed lines depicting the 3rd order polynomials fits used to derive the straight-line continuum points, grey offset (-0.01) dashed lines depicting the 3rd order polynomial fits used to derive band minima and black dotted lines depicting the straight-line apparent continua that will be used to isolate the features of interest for band centring and area derivation. **Note:** These are low resolution versions of the figures.

### 3.4. Band area

Is a measure of the area of a band which can be derived in a number of different, but complementary, methods, and expressed as a dimensionless value which will always require some definition based on the function(s) used for area derivation (see [Appendices A and B](#)). Band area should ideally be a measure of the area of the absorption bound by some form of continuum removal which has isolated the absorption, and not the area of a Gaussian or other curve used for fitting or modelling. If the band areas calculated and shared in a manuscript are based on the areas of a number of fit Gaussians, the definitions provided in the text for band areas must include a description of the derivation of the fit curves which specifically notes how the full area was calculated.

All of the possible derived metrics are of potential importance, but we consider band minima the requisite principal metric and the focus of the methodology we are espousing. We would suggest band centre, depth and area follow in importance and form a set of four primary metrics that should be referred to when discussing spectra, and those for which the methodologies for derivation should be robust, absolutely reproducible, easy to understand and completely transparent (i.e., not buried in the 'black-box' of proprietary, complex, or expensive software).

#### 3.4.1. Band minima

The minimum wavelength position for any absorption or band minima is the most universally applicable metric as it is measured from existing spectra unaltered by removal of an imposed or apparent continuum. We advocate for the use of 3rd order polynomial functions (as this is the historical norm) to derive the band minima of the spectral absorptions of interest because it is optimally suited for a symmetric curve and will not over-fit the data. [Fig. 1](#) is an example of a laboratory spectrum and, as such, it has an extremely high spectral resolution and high signal-to-noise ratio, so if one picked the lowest recorded reflectance value within this absorption, it may well be the actual minima. However, if the spectrum is of low spectral resolution, such as many observational data for solar system bodies, the real minimum value may well fall between two measured values. In order to compare the high resolution data with the low resolution data in a tenable manner, we curve fit systematically using the same methodology for both sets of spectra such that our comparisons are as robust as possible. The minima and/or centre can then be reported in two different manners, either as the actual measured minimum reflectance value nearest the infinitely resolute interpolated minimum produced by the polynomial fit or the minimum interpolated reflectance that the fit polynomial subtends. Though we advocate for the exclusive use of the nearest actual measured value, if it is to be called a minima or centre, one needs to be sure when citing previous work, whether the minima or centre reported is for an actual measured value nearest the fit, infinitely resolute theoretical minimum, or if it is an interpolated value.

The use of a polynomial and the order of the polynomial used are exceedingly important. A polynomial is used as it is a simple function and given a fixed set of coefficients, will always return a single solution. When discussing modelling (see [Appendix A and B](#)) one encounters statistical measures of goodness of fit, fit quality, or measures of fit error. Fitting a polynomial is not modelling as we would define it, and we would suggest that it is not necessary to aim for a specific measure of fit quality (though such measures are included in the [Appendix A](#) examples and should always be included with fit metrics). Rather, when curve fitting a spectrum in a repeatable, robust and transparent manner using a function that has a single solution with no free variables (e.g., a polynomial), the problem of fit quality is ameliorated when the polynomial function and full fit are reported and/or the requisite spectrum and listing of nodes are included so it can be reproduced. If one strays from the caveats regarding asymmetries about the minimum/maximum or centre wavelength noted in the

[Appendices](#), they may run into a problem, but as long as the process remains transparent, another researcher can revisit the fits.

We would suggest, even for the highest-resolution laboratory data, cubic or 3rd order polynomials should be used exclusively. Other orders of polynomials have been supplied in the [Appendices](#) for other purposes, but their use should be limited to very rare circumstances. A 2nd or 3rd order polynomial is the best compromise for the typical spectral resolutions returned by the remote and telescopic platforms characteristic of planetary science and will return consistent results across a multitude of datasets of varying resolution. Higher order polynomials have their purposes, but will 'follow the line' to too great an extent when spectra is of low resolution or noisy. It should also be noted that a 3rd order polynomial is a simple function, whose use should be limited to concave or convex portions of a curve and not both, as a curve that both rises and falls across the number of points fed into a fit of a 3rd order polynomial will skew the centre or minima values and the inflection points calculated will fall between the desired, and arguably correct values (see [Appendix A](#)).

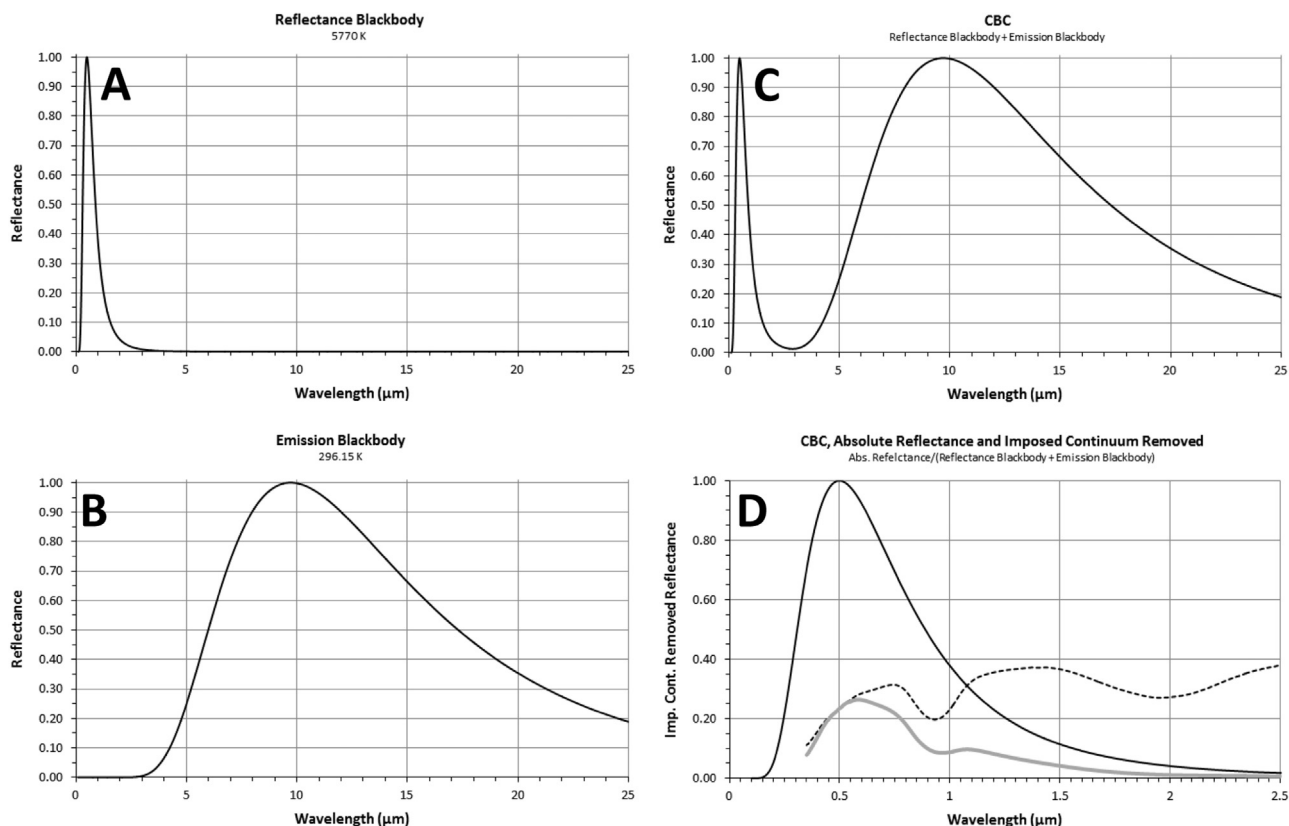
#### 3.4.2. The continuum

To derive the other primary metrics, excepting minima, one must be familiar with the concept of the continuum and how it might be applied to a spectrum. The continuum and removal of the apparent or imposed continuum by various curves or straight lines can be difficult to comprehend. However, continuum removal has but two purposes: to remove some effect imposed by the measurement/calibration process used when collecting a spectrum (imposed continuum removal); or to isolate a portion of the spectrum (apparent continuum removal).

The continuum can be thought of as a combination of curves, chiefly, the emission curves of the light source and sample, and the continuum of absorptions caused by crystal field absorptions, charge transfers, vibrational modes, overtones of vibrational modes, wings of absorption bands, optical effects, wavelength dependent scattering, and so on (see [Morris et al., 1982](#); [Clark and Roush, 1984](#)). In normal practice, as part of the calibration of the reflectance instrument, a large portion of the continuum, which is represented by the Combined reflectance and emission Blackbody Curves (CBC) of the light source and the sample, will be removed by division. As this type of continuum removal is imposed on the reflectance spectrum by the normal calibration procedures for a reflectance instrument, we refer to it as imposed continuum removal, and when we refer to spectra as reflectance spectra, imposed continuum removal is implied.

To begin, one has to understand the measurement and white referencing processes used by a particular spectrometer to collect a spectrum. For a spectrometer collecting ultraviolet thru near-infrared reflectance spectra for example, the spectra will be collected relative to a white reference reflectance standard, normally halon or Spectralon® (see [Ruff et al., 1997](#) for an infrared emission example). A spectrometer has an inherent response function (a curve or a series of curves) due to the response functions of the individual detector(s) used, the light source will also have an inherent emission curve, and so will the sample. The process of measuring relative to a standard, such as Spectralon®, creates a translation curve that will remove both the instrument's detector response function(s) and the CBC by translating the combined curves to unity, or 100% reflectance, relative to the measured unity value as defined by the white reference standard. A consequence of the process of measuring reflectance spectra in this manner is that all reflectance spectra output by the instrument will be imposed continuum removed.

[Fig. 2](#) illustrates the two components of the CBC (2A and 2B), their combination into a single curve to facilitate the removal of the continuum imposed by the measurement process, (2C) the reflectance or imposed continuum removed spectra, and (2D) the spectrum with its imposed continuum re-applied. This applies to reflectance spectra collected in the laboratory as well as remotely sensed reflectance spectra. Often asteroid spectra, for example, are referred to as I/F or



**Fig. 2.** Construction of the Combined Blackbody Curves (CBC) for a sample of pyroxene illustrating the removal of the continuum imposed on the sample spectrum by the reflectance instrument calibration/reflectance spectrum measurement process. 2A is the solar black body emission curve/reflectance blackbody curve at a temperature of 5770 K. 2B is the emission blackbody curve for the < 45  $\mu\text{m}$  powdered pyroxene sample at a temperature of 296.15 K (23  $^{\circ}\text{C}$ ). 2C is the CBC which is removed via multiplication in panel 2D to illustrate the removal of the continuum imposed on the sample by the spectral measurement process. The black dashed line is the spectrum of a pyroxene depicting the normal output of a reflectance spectrometer where the pyroxene spectra is measured relative to a white reference standard, the solid black curve is the CBC for the sample and the thicker grey line is the real spectrum of the pyroxene with its imposed continuum removed.

Intensity-over-Flux spectra, and they undergo the same process where the continuum of the light source, i.e., its emission curve, or flux is removed by dividing the returned signal, Intensity, by Flux. I/F spectra are collected in the same manner as laboratory spectra with two significant exceptions. The majority of asteroid spectra are collected by Earth-based telescopes, and a portion of the instrument calibration procedure involves accounting for, and removal of, the absorptions by Earth's atmosphere. Unlike laboratory spectra, a CBC is not used, rather the only the reflectance emission, or solar curve in Fig. 2A is used to create the imposed continuum removed spectra of 2D, and this has implications that are discussed further in the [Appendices](#).

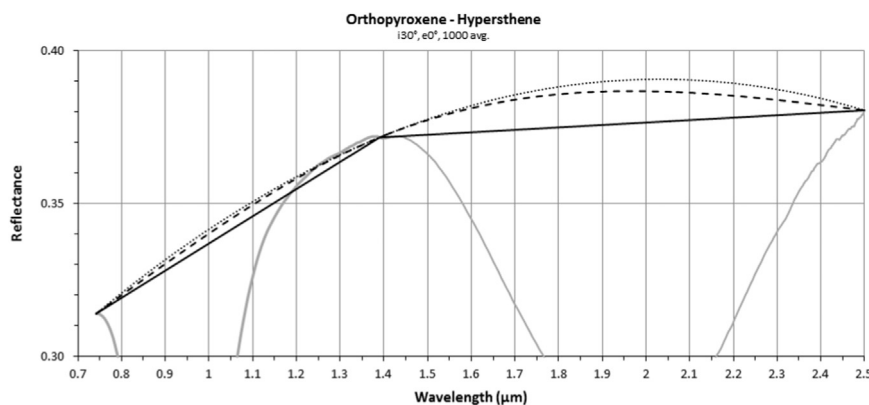
Using the CBC illustrated in 2C to remove the imposed continuum, one arrives at the true measured spectrum of the sample relative to that sample's real continuum, e.g., the thick grey spectrum in 2D. The reapplication of the imposed continuum removed via the reflectance instrument calibration/reflectance measurement process using the CBC in this manner, provided the CBC used is the same shape as the CBC initially applied to the spectrum and it is removed in the same multiplicative or additive manner, can recover the true band centres of absorptions by removing the apparent skew imparted on said absorptions by the reflectance instrument calibration/reflectance measurement procedure (see [Clark and Roush, 1984](#); [Clark, 1999](#); and the examples in [Appendix A](#)). If any modelling of spectral absorptions is to be performed, it must be completed with a spectrum where removal of the continuum imposed by the spectral measurement process has been performed.

To derive any metric other than the band minima from reflectance spectra one must somehow deal with the remaining continuum, and further remove a portion of the apparent continuum to isolate a particular spectral absorption. When an absorption band is thus

isolated, properties such as the band centre, band depth (relative to the apparent continuum) and band area can be calculated, subject to whatever constraints are imposed by the choice of apparent continuum removal method. One could choose not to remove any portion of the continuum, as is done specifically for the derivation of band minima, and still calculate a band depth. However, outside of modelling absorption features (which is further discussed in the [Appendices](#)), there are reasons for removing a portion of the apparent continuum: for isolation for defining an area, centring to remove a modicum of asymmetry, adding contrast for ease of centre wavelength determination, and to facilitate broad comparison with spectra where similar apparent continuum removal has been performed (e.g., [Clark et al., 2003](#)).

One can remove apparent continua in a number of ways and we are advocating for straight-line apparent continuum removal, i.e., the fitting of a straight line from the points of band extinction short- and long-ward of the absorption band of interest (dotted straight lines in [Fig. 1](#), solid straight lines in [Fig. 3](#)). This is the simplest method mathematically and is repeatable.

The fitting of an apparent continuum would ideally be based on sound scientific justification for the use of a particular apparent continuum removal technique and only in rare instances will straight lines or the commonly used curves such as natural cubic splines or low order moving average polynomials approximate the shape of the blackbody curves ([Fig. 3](#)). We believe the soundest scientific argument for a favoured line shape for apparent continuum removal would be a comparison of that lines shape with the shape of the CBC across the wavelength range in question. Only in very rare instances will any of the commonly use apparent continuum removal line shapes approximate the line shape of the local CBC/blackbody curve, and rarely will



**Fig. 3.** The spectrum of pyroxene with Bands I and II isolated, but uncentred, based on the derived pin points presented in Fig. 1, with fit straight-line apparent continua (solid black lines), and two oft used curved apparent continua, a 2nd order polynomial (dotted curve), and a natural cubic spline (the dashed curve). Any of the depicted straight or curved apparent continua could be used to further remove continuum from the spectrum, isolating the absorptions while centring them for derivation of area and simplification of band depth calculation. We advocate for the use of straight-line apparent continuum removal in wavelength space as demonstrated. See Appendix A for further comparison of various straight and curved line continua apparent removal methods.

using a cubic spline or low order moving average polynomial produce a different centre than a straight line. No compelling arguments have been advanced for fitting a curve in place of a straight line for apparent continuum removal. Further, the fitting of curved lines to facilitate apparent continuum removal adds significant and often unnecessary complexity to the calculation and production of continuum removed spectra. As such, we suggest the simplest and most easily repeatable option is the straight-line apparent continuum removal method, in wavelength space we advocate (see the curved continua comparison in Appendix A and discussion in Appendix B) which also allows for the widest comparison with previous studies performed where straight-line apparent continuum removal was applied.

### 3.4.3. Pin points

Picking the points to which the straight-line continuum is pinned depends most on the goal one hopes to achieve or the method chosen, e.g., while fitting the example pyroxene spectrum, we have used 3rd order polynomials to derive the inflection points/extinction points/pin points for the areas of absorption features in the wavelength regions of interest (see Fig. 1), though there are other methods for deriving pin points. One common procedure is to use the tangent points associated with straight-line continuum fit. For example, if one were working with a set of spectra, for instance, a set of silica samples as in Rice et al. (2013), or Milliken and Mustard (2005), or a set of pyroxene samples of varying grains sizes, as in Craig et al. (2008), or a set of spectra of mixtures of pyroxene and olivine, as in Cloutis et al. (1986), one might decide on a fixed set of pin points which always remain inside of the extinction points between adjacent bands for an entire set of samples and ignore the changes in extinction/pin point values from sample to sample. It may sometimes be preferable to perform the centring in the same manner as previously fit data. When the pin points are known, or the methods used are known, such as using fixed pin points at 0.7 and 2.4  $\mu\text{m}$ , and using a derived pin point for the point of extinction between Bands I and II as was done in Cloutis et al. (1986), deriving pin points using second order polynomials and using a fixed 2.5  $\mu\text{m}$  pin point if an extinction between 2 and 2.5  $\mu\text{m}$  was not obvious as in Storm et al. (2007), or using the two peaks on either side of the absorption, defined in an unknown manner, with a fixed 2.5  $\mu\text{m}$  pin point as in Dunn et al. (2010) and (2013), those methodologies may be preferred.

For the sake of consistency across data sets we would suggest that deriving the continuum pin points using 3rd order polynomial fitting should be standard practice, unless the pin points cannot be calculated, and must be picked. When fitting pin points, one should adhere to the same set of simple constraints suggested for deriving band minima and centre polynomial fits, and similarly for the fitting of minima or centres

(see Appendix A).

The derivation of pin points can be problematic when there is no obvious point of band extinction, i.e., no obvious region of convexity to fit with a polynomial function. In our pyroxene example, we have chosen to pin the long wavelength end of the straight-line continuum between 1.39 and 2.5  $\mu\text{m}$  to the measured reflectance value at 2.5  $\mu\text{m}$ . This just happens to be common practice when fitting near-infrared spectra for asteroid analog materials measured in the laboratory due to detector limitations and telescopic asteroid spectra which also commonly end at or about 2.5  $\mu\text{m}$  due to atmospheric water vapour issues (e.g., Cloutis et al., 1986; Storm et al., 2007; Dunn et al., 2010, 2013). In this instance, it is a fortunate coincidence, but there are no existing selection rules for creating a pin point if an obvious extinction point does not exist. What is most important, as with all other curve fitting, is that one share the positions of the pin point(s), the methods and if appropriate, the reasoning.

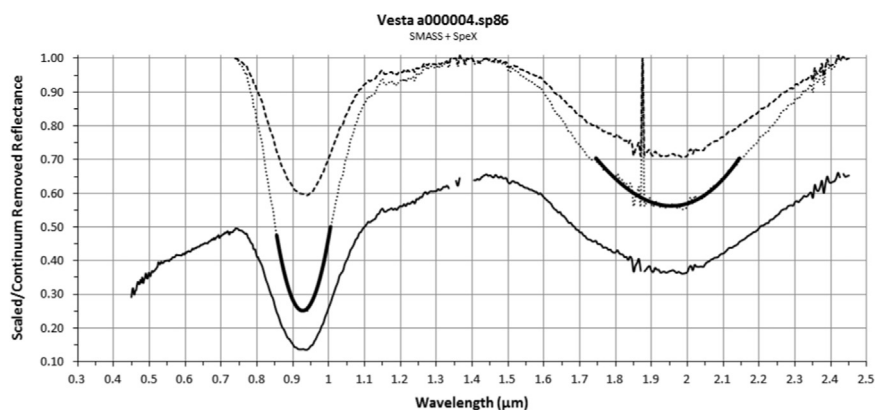
### 3.4.4. Straight-line apparent continuum removal

With pin points derived or chosen, the straight line segments for apparent continuum removal are produced, which are most often referred to as straight-line continua, and the straight line apparent continuum is mathematically removed so band depth for the continuum removed band, band centres and band areas can be calculated. The straight line continuum can be removed by dividing it out, i.e.,  $\text{sample}_y \div \text{continuum}_y$  for a given  $x$ , which centres the absorption, and/or it could be removed by linear translation, i.e.,  $(1 - \text{continuum}_y) + \text{sample}_y$  for a given  $x$ . The method of linear translation to unity is useful for visualization and for calculation of area without the contrast stretching that occurs when centring absorptions via division.

Both methods are illustrated in Fig. 4, where the solid black spectrum is the SMASS+SpeX (Rayner et al., 2003) scaled spectrum of Vesta, the dotted line spectrum is Bands I and II centred and the dashed line spectrum is Band I and II translated linearly to unity.

### 3.4.5. Band centre

Band centres are fit in the same manner as band minima, using 3rd order polynomials, but fit to spectra where the band has been centred. It is important that centres only be reported when an apparent continuum has been removed by division and the spectra have experienced the contrast stretch that is applied by this method of apparent continuum removal. It must also have been performed on spectra where the continuum imposed by the reflectance measurement process is intact, and a straight or curved line apparent continuum has been removed to isolate the absorption in question. It may be tempting to use the term band centre to denote any measure of a minimum on a band that has had the continuum altered in some way or a portion



**Fig. 4.** Straight-line apparent continuum removed spectra of the two isolated bands of interest for asteroid (4) Vesta. The dotted spectrum has been straight line apparent continuum removed via division, i.e., centred, and the dashed line spectrum has been straight line apparent continuum removed via linear translation. The solid black spectrum is the unaltered (but offset) SMASS+SpeX spectra of Vesta. The heavy black lines which overlay the absorption minima for the centred dotted spectrum depict the polynomials fit to derive the band centres. The spikes in the two continuum removed spectra within Band II, and the portions of the continuum removed spectra which are above unity are discussed, as is the fitting of data with gaps, in more detail in the [Appendices](#).

removed as in infrared emission spectra, but, given the ubiquity of its use for near-infrared reflectance spectra, the term band centre should be reserved for near-infrared use.

The band centre value is derived by fitting a 3rd order polynomial function to a subsection of the centred spectrum (solid black fits in [Fig. 4](#)). Best practice is to pick a portion of the centred spectrum which is reasonably symmetric, visually, about the minimum value with approximately the same number of points either side of the approximate centre. One can stray marginally from an area of symmetry around the approximate centre value, but with significant skew, the derived centre value will be inappropriate (see [Appendices A and B](#) for further examples and explanation).

#### 3.4.6. Band depth

Band depth can be calculated in a number of different manners depending on the spectrum under investigation. Regardless, band depth is always the depth to the minimum or centre along a straight line, which intersects the band minima (though not necessarily a fit minima), or centre, tangent to a fit straight-line continuum, unity, or following from some other defined position in the parameterized spectrum. Unlike centres, band depth has no standardized meaning and this requires that it be defined in each instance (see [Rice et al., 2013](#), for a thoroughly documented example of the methods suggested by [Clark and Roush, 1984](#)).

#### 3.4.7. Band area

One can calculate the area of a band using one of the three methods we have provided: using chords, the trapezoid rule or integration of regression, or utilize some other method. What matters in the calculation of area is not the method so much as consistent application of one method across the spectra being fit or a method that is consistent with the spectra to which they will be compared. Using all three methods will result in different measured values for area (which can be made to reflect real numbers) that when ratioed will return very nearly the same result (see [Appendix A](#) for a comparison). When representing Band Area Ratio (BAR), it is of utmost importance that BAR be calculated using area measures derived from bands centred by dividing out of straight-line apparent continua. When centring a band, one is dividing out the apparent continuum which can be described as linearly moving the numerically highest pin point reflectance value to unity, while swinging the rest of the values up such that the straight-line apparent continuum becomes a straight horizontal line at unity. This results in a contrast stretch, i.e., the depth of the band increases, and the degree to which it deepens is governed by the degree to which the straight-line continuum has swung to achieve unity.

Referencing the straight dotted lines in [Fig. 1](#), the straight-line apparent continuum running from 0.742 to 1.39  $\mu\text{m}$  has significantly more positive slope than the straight line running from 1.39 to 2.5  $\mu\text{m}$ . When centring both bands, the contrast stretch factors (a number calculated by ratioing the depth of the centre of the centred band to the depth of that same wavelength value for the same band translated linearly to unity) are 3.02 and 2.66 for Bands I and II respectively. This means that the bands have not been stretched by the same factor when centering, and the BAR is inherently altered as a result, e.g., the BAR of the centred bands is 1.93, while the BAR for the uncentred bands is 2.20, so, if one were plotting the band area ratios on the S Asteroid subtype plot of Cloutis and Gaffey ([Gaffey, 1993](#)), not recognizing the difference could lead to significant misinterpretation. The issue is readily apparent graphically in the centred versus uncentred band comparison in [Fig. 4](#).

The spectrum of (4) Vesta, [Fig. 4](#), illustrates another set of problems that are commonly encountered when dealing with remotely obtained spectra, which are rarely encountered in laboratory spectra: noise and missing data. Noisy data are the primary reason for using a 3rd order polynomial for fitting, as 3rd order polynomial fits are generally insensitive to the high-frequency noise encountered in spectral data sets (for modelling another choice may be appropriate), but the noise can present another problem. When fitting the curves for the apparent continuum pin points, the noise may affect the positioning and may result in portions of the continuum removed spectrum exceeding a value of 1, or unity. In these instances, the calculations for area we have devised will subtract the calculated areal values which surpass unity, or 100% reflectance. This may also occur in high resolution, relatively noise free spectra, where centred reflectance values may exceed unity after apparent continuum removal. In these instances, the calculations for area treat the values exceeding unity in a manner defined by the chosen areal calculation method (see [Appendix A](#) for illustrative examples).

When dealing with missing data, which is also illustrated in [Fig. 4](#), it is up to the person performing the fitting what sort of gap size is tolerable, and how the data gap may be treated. When data gaps are small, i.e., on the order of a few nanometres, for the purposes of fitting, we have provided polynomial fitting routines that can fit data with these small gaps and unique functions for centred and linear translation apparent continuum removal. When removing the apparent continuum for gapped data, the functions replace all missing y values with 1's to provide a visual reminder that one will have to use a function specific to gapped data when calculating band areas. Like the polynomials specific to gaps, idealized functions have been created to ignore the data gaps for the three areal measurement methods



provided (see [Appendices](#)). We suggest that ignoring the gaps for area calculation is the mathematically soundest method when data gaps are small and fully removing gaps and allowing the trapezoidal area function to in-fill them in a linear manner is most efficacious when gaps are minor. However, for instances where gaps are quite large one may be better served by in-filling missing data points using polynomial interpolation. Functions for doing so are provided in the [Appendices](#) but suggesting any sort of framework for their use is exceedingly difficult; deciding when and how to interpolate missing data points should be made on a case by case basis. In the example provided in [Fig. 4](#), regarding the SMASS + SpeX spectrum of Vesta, the BAR which results from areas calculated using the trapezoidal area with gapped values being 1 (which results in no areal values being reported for gaps, and gap adjacent areas) is 1.39, versus BAR values of 1.41 calculated using trapezoid functions for the same spectrum where gaps were in-filled using polynomial and/or linear interpolation functions. If one plots these on the Cloutis/Gaffey S-Asteroid plot of [Gaffey \(1993\)](#), or the modified plot of [Dunn et al. \(2013\)](#), it could be argued that the effect in this instance is negligible, but the matter is debatable, as in cases where a sample's Band I centre versus BAR might have it subtend an S asteroid subtype boundary, what was a negligible difference in BAR value may now prove quite significant. What matters is not necessarily how it was done, but that the methods are disseminated such that cross comparisons with other data sets, or comparisons by other researchers, are valid.

#### 4. Modelling

Modelling spectra deserves particular attention as it is an oft misunderstood term which can lead to confusion; we propose that the term “modelling,” in the context of spectral absorptions, be restricted to a narrow definition. We would suggest that any model must be modelling the real physical processes responsible for the creation of an absorption.

We suggest this constriction on the bounds of the use of the word ‘model’ as the standard procedure with ultraviolet through far-infrared spectra, after the derivation of band minima, would be to isolate an absorption by removing a straight-line apparent continuum for further derivation of spectral parameters. This is normally performed on a reflectance spectrum that has its imposed continuum resulting from the measurement process intact, which means that the spectrum one is attempting to model has absorption shapes skewed by the reflectance instrument calibration procedure, and then those same absorptions are skewed further via removal of more of the continuum in a rather arbitrary fashion. Procedurally, this is acceptable, when both the researcher and the intended audience understand that the process of measurement has removed the reflectance continuum and the absorption in question has been centred as per standard practice and definition.

Where many seem to go awry is in assuming that any fitting of a spectrum where any portion of the continuum has been twice altered is in some way modelling of the spectrum. As noted earlier, any altering of the spectrum, by application of an imposed continuum through measurement processes or via removal of straight or curved line apparent continua either by division or translation will alter the spectrum such that the positions and shapes of absorptions have no basis in physical reality, meaning that no actual modelling, as we define it, can occur.

The pyroxene spectrum of [Fig. 5](#) is an example of modelling of the real spectrum, with its imposed continuum in wavelength space removed (as seen in [Fig. 2D](#)), modelled in energy space using summed Gaussians, where they are modelling the physical absorptions creating the spectrum and are as representative as they can be of the full spectrum, given that the Gaussians used are there purely to model crystal field transition and are ignoring other processes influencing the spectrum. For this example, 6 Gaussians are used, although only three

Gaussian centre positions are of importance for the pyroxene spectrum, i.e., those centred at 0.92, 1.105 and 1.965  $\mu\text{m}$  (1.345, 1.121 and 0.632 eV). Three of the Gaussians used model the spectral features due to  $\text{Fe}^{2+}$  in the M1 and M2 coordination sites in the pyroxene and the other three serve as placeholders for a number of absorptions, and important to note, while only three Gaussian centres are of importance, the 6 modelled absorptions, and an additional 7 would be required to thoroughly model this portion of the pyroxene samples spectrum. This model shown in [Fig. 5](#) was produced via a combination of hand and automated fit optimization, and the residual is offset plus 0.1 above unity for clarity. One can see a representation of the quality of the model via the general lack of deviation about 1.1 of the residual line (RMS 0.004). This qualifies as modelling, by our definition, as it is based in the fundamental physics that one would use to derive the idealized energy values for the crystal field transitions and overtones responsible for the combination absorptions which produce the large features of interest in the spectrum.

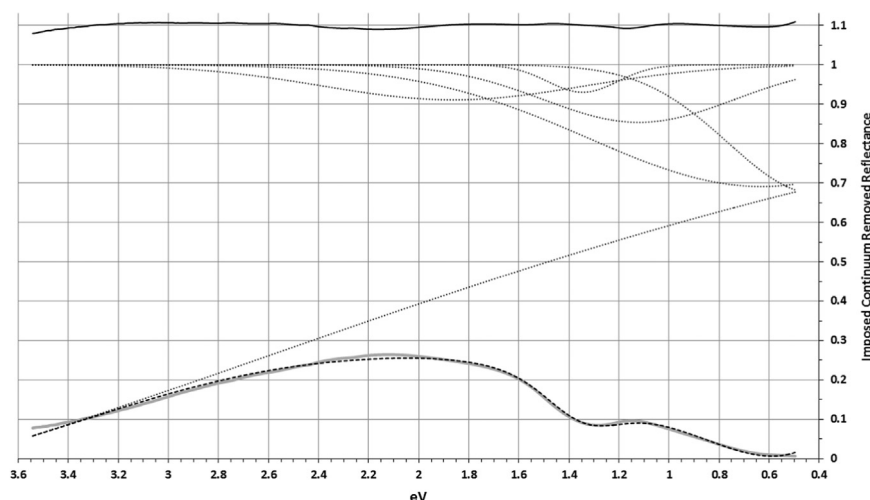
The model is illustrative of two significant points: 1) while the fit residual is low, we have not adequately modelled all 13 of the absorptions that are required at a minimum in this wavelength range to thoroughly model the spectrum (two of which will be placeholders for absorptions we cannot see, but are quite powerful, on either end of the spectral range). As such, we cannot be fully confident that the Gaussian centres, depths and FWHM are accurate models of reality, as they are based on the energy centres of real modelled and measured absorptions with sound physical explanations (e.g., [Rossman, 1980](#); [Burns, 1993](#)); and, 2) despite using Gaussians whose centres are based on the physics of absorptions in pyroxene, a low residual model can be produced with Gaussians centred at disparate locations which have no associated physical process and still produce an equally low residual (see examples in [Appendix A](#)). The fine dotted lines in [Fig. 5](#) are the individual Gaussians used and minor changes in one modelled absorption can largely be compensated for by a minor change in another, such that one could easily shift centre positions, FWHM values and scaling values of the Gaussians to produce a low residual model but those values would not be physically meaningful. Akin to MGM, where some set of reasonably fixed starting parameters are required, this sort of brute force modelling requires a set of starting values that are based on the physics of the absorptions to be valid and valuable.

Further problems with the modelling exist because the removal of the continuum imposed by the reflectance measurement process is not absolutely correct, given that the applied solar emission curve/reflectance continuum may not adequately replicate the combined effect of the light source and instrument correction curves applied to the spectrum being measured and the emission continuum may not reflect the sample temperature adequately (as the interacting volume of the sample under measurement may be heated above the ambient room temperature by the light source), such that we may not know the exact values for the necessary curves with absolute certainty.

In general, one should approach modelling of any spectra with caution as, even with monomineralic samples where the effects of grain size, packing, porosity, temperature and so on are known, and may be quantifiable individually, we cannot yet adequately constrain the effects each, in combination, is having on a particular sample's spectrum.

When solid samples of known mineralogy are being modelled, we will always have issues when modelling that are difficult to overcome which arise from known complications such as minute errors in the curves we are using for imposed continuum removal and re-application, not being able to adequately account for wavelength dependent scattering, optical effects, overtones, combination bands, hot bands, Fermi resonances, etc., and, extra contributors in samples from adsorbed species, site substitutions, inclusions and so on.

When samples are mixed assemblages of minerals of different



**Fig. 5.** Gaussian modelling of the pyroxene sample in eV space, with the imposed continuum removed in wavelength space such that the Gaussians used model the real absorptions responsible for the majority of the spectrum. Dotted lines depict the 6 Gaussians used, the thick grey spectrum is of the pyroxene sample with its imposed continuum removed, and the dashed black line residing within the thick grey spectrum is the summed Gaussian model. The quality of the fit is measured by the residual, where the better the fit, the less the residual varies about the 1.1 line (RMS 0.004).

grains sizes in the laboratory, mixed assemblages with different packing densities, grain sizes, and so on, or natural rocks, issues simply compound and physically meaningful spectral modelling may well be impossible.

## 5. Further parameterization and manipulation of spectra

Beyond methods for the derivation of the four principal empirical curve fitting metrics, i.e., band minima, centre, depth, and area, we have included several other metrics, methods for their derivation, and illustrative examples of common functions performed on spectra that deserve further explanation and/or those that are almost universally harmful and benefit from graphical explanations. Chief among these are a variety of conversions, interpolation methods, measures of skew, curved continuum removal, spectral convolution, automated and manual curve deconvolution/Gaussian fit optimization methods and modelling comparisons, further derivation of the line shapes for Lorentzian and Pseudo-Voigt curves and an exploration of smoothing functions, all of which can be found in the [Appendices](#).

## 6. Summary and conclusions

This work presents fitting routines written for Microsoft Excel® that allow one to empirically curve fit spectra using polynomial fits of raw, or straight-line apparent continuum removed spectra to mathematically and repeatably define band minima or centres with linear least squares, best-fit polynomials, in a robust and transparent cut-and-paste manner. The included functions do not rely on macros or any other add-ins to Excel that can malfunction if the included workbook or workbooks derived from it are shared across Excel platforms.

We advocate for the use of 3rd order polynomials almost exclusively for the derivation of band minima and centres, as a cubic polynomial has been found through extensive experimentation to be most efficacious across data sets of varying spectral resolution and signal-to-noise ratios. We suggest that straight-line apparent continuum removal in wavelength space be the exclusive method for centring as-measured, ultraviolet through near-infrared absorption bands, and that the term's common definition change to reflect both this, and the exclusion of the use of curved continua for band centring. We hope that the confusion that arises when the terms fitting and modelling are used interchangeably can be ameliorated by respecting the definitions we adopt herein, where modelling is only used to denote spectra modelled in energy space with real continua intact/imposed continua removed. We also

expect that, following a thorough review of the material included in the [Appendices](#), it will be abundantly clear why fully parameterized fits of spectra need to be included in or alongside published manuscripts if the planetary science community is to gain from the publication of manuscripts including and/or referencing spectra. One will find on review, that most studies which perform empirical curve fitting or curve deconvolution do not include enough data about the curve fitting or deconvolution, or parameterization thereof, to allow one to repeat the fitting for the spectra referenced within.

In an ideal world, we would hope that all researchers would always be on the same page regarding the use of one universally applicable succinct curve fitting methodology, where either the curve fits themselves are always shared or the methods are absolutely transparent and all of the requisite nodes are included so the fitting is repeatable. For those well-versed in the collection and application of spectroscopic data, the rationale for curve fitting is abundantly clear, but how to curve fit is often times not. We suggest this needs to change with the application of a consistent, robust, repeatable and transparent methodology, where spectra with included fits are shared openly, or spectra with complete sets of nodes/parameters/fit statistics are included in any and all manuscripts that reference a fit metric, such that the fit metrics can be referenced in perpetuity without questions regarding their validity. Or, spectra should always be included so each researcher could then later apply their favoured curve fitting methodology to every spectrum to which their research might refer. The former suggestion is significantly less time consuming and easier to achieve.

## Acknowledgements

MAC gratefully acknowledges funding from the Natural Sciences and Engineering Research Council of Canada (NSERC), the NSERC Collaborative Research and Training Experience Program-Technologies and Techniques for Earth and Space Exploration, and the Ontario Graduate Scholarship Program that contributed to this work. The HOSERLab at the University of Winnipeg, where a portion of the presented spectra were collected, was established with funding provided by the Canadian Foundation for Innovation, the Manitoba Research Innovations Fund and the Canadian Space Agency. A portion of the data utilized in this publication and its appendices was obtained from, and made available by The MIT-UH-IRTF Joint Campaign for NEO Reconnaissance. The IRTF is operated by the University of Hawai'i under cooperative agreement no. NCC 5-538 with the National Aeronautics and Space Administration, Office of Space

Science, Planetary Astronomy Program. The MIT component of this work is supported by NASA grant 09-NEOO009-001, and by the National Science Foundation under Grants Nos. 0506716 and 0907766.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2016.11.018.

## References

- Abe, M., Takagi, Y., Kitazato, K., Abe, S., Hiroi, T., Vilas, F., Clark, B.E., Abell, P.A., Lederer, S.M., Jarvis, K.S., Nimura, T., Ueda, Y., Fujiwara, A., 2006. Asteroid Itokawa from the Hayabusa spacecraft. *Science* 312, 1334–1338.
- Abell, P.A., Vilas, F., Jarvis, K.S., Gaffey, M.J., Kelley, M.S., 2007. Mineralogical composition of (25143) Itokawa 1998 SF36 from visible and near-infrared reflectance spectroscopy: evidence for partial melting. *Meteorit. Planet. Sci.* 42 (12), 2165–2177.
- Bakker, W.H., van Ruitenbeek, F.J.A., van der Werff, H.M.A., Zegers, T.E., Oosthoek, J.H.P., Marsh, S.H., van der Meer, F.D., 2014. Processing OMEGA/mars express hyperspectral imagery from radiance-at-sensor to surface reflectance. *Planet. Space Sci.* 90, 1–9.
- Berger, J.A., King, P.L., Green, A., Craig, M.A., Spilde, M.N., Wright, S.P., Kunkel, T.S., Lee, R.J., 2015. Effect of halite coatings on thermal infrared spectra. *J. Geophys. Res.: Solid Earth* 120, 2162–2178.
- Binzel, R.P., Rivkin, A.S., Bus, S.J., Sunshine, J.M., Burbine, T.H., 2001. MUSES-C target asteroid (25143) 1998 SF36: a reddened ordinary chondrite. *Meteorit. Planet. Sci.* 36, 1167–1172.
- Burbine, T.H., Buchanan, P.C., Dolcar, T., Binzel, R.P., 2009. Pyroxene mineralogies on near-Earth vestoids. *Meteorit. Planet. Sci.* 44 (Nr 9), 1331–1341.
- Burns, R.G., 1993. *Mineralogical Applications of Crystal Field Theory*. Cambridge University Press, Cambridge, 551.
- Buz, J., Ehlmann, B.L., 2014. Effects of Grain Size on the Reflectance Spectroscopy of Olivine in the Vis-nir and the Derivation of Olivine Composition Using Modified Gaussian Model. LPSC XLV, abstract #2810.
- Clark, B.E., Binzel, R.P., Howell, E.S., Cloutis, E.A., Ockert-Bell, M., Christensen, P., Barucci, M.A., DeMeo, F., Lauretta, D.S., Connolly, H., Jr., Soderberg, A., Hergenrother, C., Lim, L., Emery, J., Mueller, M., 2011. Asteroid (101955) 1999 RQ36: spectroscopy from 0.4 to 2.4  $\mu\text{m}$  and meteorite analogs. *Icarus* 216, 462–475.
- Clark, R.N., 1980. A large-scale interactive one-dimensional array processing system. *Publ. Astron. Soc. Pac.* 92, 221–224.
- Clark, R.N., 1981. Water frost and ice: the near-infrared spectral reflectance 0.65–2.5  $\mu\text{m}$ . *J. Geophys. Res.* 86 (B4), 3087–3096.
- Clark, R.N., 1993. *SPECTRUM Processing Routines User's Manual Version 3 (program SPECPR)*: U.S. Geological Survey, Open-File Report 93-595.
- Clark, R.N., 1999. Spectroscopy of rocks and minerals and principles of spectroscopy. In: Renz, A.N. (Ed.), *Manual of Remote Sensing*. John Wiley, New York, 3–58.
- Clark, R.N., Roush, T.L., 1984. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.* 89 (B7), 6329–6340.
- Clark, R.N., Swayze, G.A., Livo, K.E., Kokaly, R.F., Sutley, S.J., Dalton, J.B., McDougal, R.R., Gent, C.A., 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res.* 108 (E12), (5-1-5-44).
- Clenet, H., Pinet, P., Daydou, Y., Heuripeau, F., Rosemberg, C., Baratoux, D., Chevrel, S., 2011. A new systematic approach using the modified Gaussian model: insight for the characterization of chemical composition of olivines, pyroxenes and olivine-pyroxene mixtures. *Icarus* 213, 404–422.
- Cloutis, E.A., Gaffey, M.J., 1991. Pyroxene spectroscopy revisited: spectral-compositional correlations and relationships to geothermometry. *J. Geophys. Res.* 96 (E5), 22809–22826.
- Cloutis, E.A., Gaffey, M.J., Jackowski, T.L., Reed, K.L., 1986. Calibrations of phase abundance, composition, and particle size distribution for olivine-orthopyroxene mixtures from reflectance spectra. *J. Geophys. Res.* 91 (B11), 11641–11653.
- Craig, M.A., Cloutis, E.A., Reddy, V., Bailey, D.T., Gaffey, M.J., 2008. The Effects of Grain Size, < 10 $\mu\text{m}$  – 4.75 mm, on the Reflectance Spectrum of Planetary Analogs From 0.35–2.5  $\mu\text{m}$ . LPSCXXXIX, abstract #2082.
- Doetsch, G. Von, 1928. The elimination of the Doppler effect in spectroscopic fine structure and exact determination of the components. *Z. Phys.* 49, 705–730.
- Dunn, T.L., McCoy, T.J., Sunshine, J.M., McSween, H.Y., Jr., 2010. A coordinated spectral, mineralogical, and compositional study of ordinary chondrites. *Icarus* 208, 789–797.
- Dunn, T.L., Burbine, T.H., Bottke, W.F., Jr., Clark, J.P., 2013. Mineralogies and source regions for the near-Earth asteroids. *Icarus* 222, 273–282.
- Farr, T.G., Bates, B.A., Ralph, R.L., Adams, J.B., 1980. Effects of overlapping optical absorption bands of pyroxene and glass on the reflectance spectra of lunar soils. In: *Proceedings of the Lunar and Planetary Science Conference XI*, pp. 719–729.
- Fiebert-Beyer, S.K., Gaffey, M.J., Kelley, M.S., Reddy, V., Reynolds, C.M., Hicks, T., 2011. The Maria asteroid family: genetic relationships and a plausible source of the mesosiderites near the 3:1 Kirkwood gap. *Icarus* 213, 524–537.
- Gaffey, M.J., 1976. Spectral reflectance characteristics of the meteorite classes. *J. Geophys. Res.* 81 (5), 905–920.
- Gaffey, M.J., 1983. The asteroid (4) Vesta: Rotational spectral variations, surface material heterogeneity, and implications for the origin of the basaltic achondrites. LPSCXIV, abstract #1118.
- Gaffey, M.J., 1993. Mineralogical variations within the S-type asteroid class. *Icarus* 106, 573–602.
- Gaffey, M.J., 1997. Surface lithologic heterogeneity of asteroid 4 Vesta. *Icarus* 127, 130–157.
- Gaffey, M.J., 2008. Interpreting asteroid spectra – avoiding the three great mistakes. In: *Asteroids, Comets, Meteors Conference*, abstract #8162.
- Gaffey, M.J., 2010. Space weathering and the interpretation of asteroid reflectance spectra. *Icarus* 209, 564–574.
- Gaffey, M.J., McCord, T.B., 1977. Asteroid surface materials: mineralogical characterizations and cosmological implications. *Proc. Lunar Sci. Conf.* 8, 113–143.
- Gaffey, M.J., McCord, T.B., 1978. Asteroid surface materials: mineralogical characterizations from reflectance spectra. *Space Sci. Rev.* 21, 555–628.
- Gaffey, M.J., Gilbert, S.L., 1998. Asteroid 6 Hebe: the probable parent body of the H-type ordinary chondrites and the IIE iron meteorites. *Meteorit. Planet. Sci.* 33, 1281–1295.
- Gaffey, M.J., King, T.V.V., Hawke, B.R., Cintala, M.J., 1982. Asteroid Spectral Variations: Implications for Composition and Surface Processes, LPSCXIII, abstract #1128.
- Gaffey, M.J., Cloutis, E.A., Kelley, M.S., Reed, K.L., 2002. Mineralogy of asteroids. In: Bottke, W.F., Jr., Cellino, A., Paolich, P., Binzel, R.P. (Eds.), *Asteroids III*. University of Arizona Press, Tucson, 183–204.
- Gallie, E.A., Lyder, D.A., Rivard, B., Cloutis, E.A., 2008. Technical Note: equivalence of modified Gaussian model (MGM) in wavenumber and Gaussian in wavelength for deconvolution of hyperspectral reflectance spectra. *Int. J. Remote Sens.* 29 (14), 4089–4096.
- Hergenrother, C.W., Nolan, M.C., Binzel, R.P., Cloutis, E.A., Barucci, M.A., Michel, P., Scheeres, D.J., Drouot d'Aubigny, C., Lazzaro, D., Pinilla-Alonso, N., Campins, H., Licandro, J., Clark, B.E., Rizk, B., Beshore, E.C., Lauretta, D.S., 2013. Lightcurve, color and phase function photometry of the OSIRIS-Rex target asteroid (101955) Bennu. *Icarus* 226, 663–670.
- Kaper, H.G., Smits, D.W., Schwarz, U., Takakubo, K., van Woerden, H., 1966. Computer analysis of ordered distributions into Gaussian components. *Bull. Astron. Inst. Neth.* 18, 465–487.
- Klima, R.L., Pieters, C.M., Dyar, M.D., 2007. Spectroscopy of synthetic Mg-Fe pyroxenes I: spin-allowed and spin-forbidden crystal field bands in the visible and near-infrared. *Meteorit. Planet. Sci.* 42 (2), 235–253.
- Klima, R.L., Dyar, M.D., Pieters, C.M., 2011. Near-infrared spectra of clinopyroxenes: effects of calcium content and crystal structure. *Meteorit. Planet. Sci.* 46 (3), 379–395.
- Kraft, M.D., Michalski, J.R., Sharp, T.G., 2003. Effects of pure silica coatings on thermal emission spectra of basaltic rocks: considerations for Martian surface mineralogy. *Geophys. Res. Lett.* 30 (24), 2288.
- Livo, K.E., Clark, R.N., 2014. *The Tetracorder user guide-version 4.4: U.S. Geological Survey, Open-File Report 2013-1300*.
- Von Lonn, E., 1932. Proof of the Uniqueness of the Decomposition of an Intensity Curve into its Components. *Z. Phys.* 75, 348–349.
- Lucey, P.G., Clark, R.N., 1985. Spectral properties of water ice and contaminants. In: Klinger, J. (Ed.), *Ices in the Solar System*. D. Reidel Publishing Company, France, 155–168.
- Makarewicz, H.D., Parente, M., Bishop, J.L., 2009. Deconvolution of VNIR spectra using Modified Gaussian Modeling (MGM) with automatic parameter initialization (API) applied to CRISM. *IEEE Whisp.* Article 5289046.
- McCord, T.B., Adams, J.B., Johnson, T.V., 1970. Asteroid Vesta: spectral reflectivity and compositional implications. *Science* 168, 1445–1447.
- McCord, T.B., Clark, R.N., Hawke, B.R., McFadden, L.A., Owensby, P.D., Pieters, C.M., Adams, J.B., 1981. Moon: near-infrared spectral reflectance, a first good look. *J. Geophys. Res.* 86 (B11), 10883–10892.
- McSween, H.Y., Jr., Grove, T.L., Wyatt, M.B., 2003. Constraints on the composition and petrogenesis of the Martian crust. *J. Geophys. Res.* 108 (E12), 5135.
- McSween, H.Y., Jr., Binzel, R.P., De Sanctis, M.C., Ammannito, E., Prettyman, T.H., Beck, A.W., Reddy, V., Le Corre, L., Gaffey, M.J., McCord, T.B., Raymond, C.A., Russell, C.T., the Dawn Science Team, 2013. Dawn; the Vesta-HED connection; and the geologic context for eucrites, diogenites and howardites. *Meteorit. Planet. Sci.* 48 (Nr. 11), 2090–2104.
- van der Meer, F.D., 2004. Analysis of spectral absorption features in hyperspectral imagery. *Int. J. Appl. Earth Obs. Geoinf.* 5, 55–56.
- Milliken, R.E., Mustard, J.F., 2005. Quantifying absolute water content of minerals using near-infrared reflectance spectroscopy. *J. Geophys. Res.* 110 (E12001), 25.
- Morris, R.V., Neely, S.C., Mendell, W.W., 1982. Application of Kubelka-Munk theory of diffuse reflectance to geologic problems: the role of scattering. *Geophys. Res. Lett.* 9 (2), 113–116.
- Nakamura, T., Noguchi, T., Tanaka, M., Zolensky, M.E., Kimura, M., Tsuchiyama, A., Nakato, A., Ogami, T., Ishida, H., Uesugi, M., Yada, T., Shirai, K., Fujimura, A., Okazaki, R., Sandford, S.A., Ishibashi, Y., Abe, M., Okada, T., Ueno, M., Mukai, T., Yoshikawa, M., Kawaguchi, J., 2011. Itokawa dust particles: a direct link between S-type asteroids and ordinary chondrites. *Science* 333, 1113–1116.
- Parente, M., Makarewicz, H.D., Bishop, J.L., 2011. Decomposition of mineral absorption bands using nonlinear least squares curve fitting: application to Martian meteorites and CRISM data. *Planet. Space Sci.* 59, 423–442.
- Ramsey, M.S., Christensen, P.R., 1998. Mineral abundance determination: quantitative deconvolution of thermal emission spectra. *J. Geophys. Res.* 103 (B1), 577–596.
- Rayner, J.T., Toomey, D.W., Onaka, P.M., Denault, A.J., Stahlberger, W.E., Vacca, W.D., Cushing, M.C., Wang, S., 2003. SpeX: a medium-resolution 0.8–5.5  $\mu\text{m}$  spectrograph and imager for the NASA infrared telescope facility. *Publ. Astron. Soc. Pac.* 115,

- 362–382.
- Reddy, V., Sanchez, J.A., Bottke, W.F., Cloutis, E.A., Izawa, M.R.M., O'Brien, D.P., Mann, P., Cuddy, M., Le Corre, L., Gaffey, M.J., Fujihara, G., 2014. Chelyabinsk meteorite explains unusual spectral properties of Baptistina asteroid family. *Icarus* 237, 116–130.
- Rice, M.S., Cloutis, E.A., Bell, J.F., III, Bish, D.L., Horgan, B.H., Mertzman, S.A., Craig, M.A., Renaut, R.W., Gautason, B., Mountain, B., 2013. Reflectance spectra diversity of silica-rich materials: sensitivity to environment and implications for detection on Mars. *Icarus* 223, 499–533.
- Rossmann, G.R., 1980. Pyroxene spectroscopy. In: *Pyroxene. Reviews in Mineralogy No. 7. Mineralogical Society of America, Washington, DC*, 93–115.
- Ruff, S.W., Christensen, P.R., Barbera, P.W., Anderson, D.L., 1997. Quantitative thermal emission spectroscopy of minerals: a laboratory technique for measurement and calibration. *J. Geophys. Res.* 102 (B7), 14889–14913.
- Singer, R.B., 1981. Near-infrared spectral reflectance of mineral mixtures: systematic combinations of pyroxenes, olivine, and iron oxides. *J. Geophys. Res.* 86 (B9), 7967–7982.
- Storm, S., Bus, S.J., Binzel, R.P., 2007. Olivine-pyroxene distribution of S-type asteroids in the main belt. *Bull. Am. Astron. Soc.* 39, 448.
- Sunshine, J.M., Pieters, C.M., 1993. Estimating modal abundances from the spectra of natural laboratory pyroxene mixtures using the modified Gaussian model. *J. Geophys. Res.* 98 (E5), 9075–9087.
- Sunshine, J.M., Pieters, C.M., 1998. Determining the composition of olivine from reflectance spectroscopy. *J. Geophys. Res.* 103 (E6), 13675–13688.
- Sunshine, J.M., Pieters, C.M., S. F. Pratt, S.F., 1988. Gaussian Analysis of Pyroxene Reflectance Spectra, LPSC XIX, Abstract #1151.
- Sunshine, J.M., Pieters, C.M., Pratt, S.F., 1989. Mathematical Deconvolution of Mineral Absorption Bands, LPSCXX, abstract #1087.
- Sunshine, J.M., Pieters, C.M., Pratt, S.F., 1990. Deconvolution of mineral absorption bands: an improved approach. *J. Geophys. Res.* 95 (B5), 6955–6966.
- van Ruitenbeek, F.J.A., Bakker, W.H., van der Werff, H.M.A., Zegers, T.E., Oosthoek, J.H.P., Omer, Z.A., Marsh, S.H., van der Meer, F.D., 2014. Mapping the wavelength position of deepest absorption features to explore mineral diversity in hyperspectral images. *Planet. Space Sci.* 101, 108–117.