A NEW MODIFIED GAUSSIAN MODEL (MGM) USING A BAYESIAN ESTIMATION APPROACH: TOWARD AUTOMATED ANALYSIS OF PLANETARY SPECTRA. S. Sugita[1], K. Nagata[1], N. Tsuboi[1,2], T. Hiroi[3], and M. Okada[1], [1]Dept. of Complexity Sci. and Eng., University of Tokyo (5-1-5 Kashiwanoha, Kashiwa, Chiba, Japan, sugita@k.u-tokyo.ac.jp)), [2]Hewlett Packard Japan, Ltd., [3]Dept. of Geol. Sci., Brown University

We derived a new modified Gaussian model (MGM) algorithm that can objectively find the optimum number of Gaussians and has little dependence on initial parameter selection. This algorithm is suitable for automated analyses of planetary reflectance spectra, particularly for Moon, which has a tremendous amount of high-resolution spectral datasets.

Planetary Spectral Analysis: Visible and near infrared (VNIR) reflectance spectroscopy is a very powerful tool to observe planetary surfaces remotely and has revealed a wide variety of information on planetary surfaces [e.g., 1]. Because silicates often exhibit overlapped absorption bands in the VNIR range, they are often deconvolved into multiple simple (e.g., Gaussian) bands for quantitative analyses. One simple approach would be to search for the optimum combination of simple bands that minimize the difference E between observed R_{obs} and synthetic R_{synth} spectra by trying every possible parameter set $\boldsymbol{\mu}=\{\mu_k\}$. For example, E is given by

$$E(\boldsymbol{\mu}) = \sum_i \left\{ \log R_{obs}(\lambda_i) - \log R_{synth}(\lambda_i, \boldsymbol{\mu}) \right\}^2 \quad (1)$$

where λ is wavelength. However, a direct search is usually impractical because it requires an astronomical number of trials.

MGM with the Steepest Descent Method: Thus, a more efficient mathematical method is necessary. In previous studies, the steepest descent method has been used frequently to resolve this problem, in which optimal parameter μ_k is sought with the following iteration:

$$\mu_k^{m+1} = \mu_k^m - \varepsilon \left. \frac{\partial E(\mu_k)}{\partial \mu_k} \right|_{\mu_k = \mu_k^m} \quad (2)$$

where m and ε are iteration number and a constant, respectively. *Sunshine et al*. [2] propose an analytical approach using the steepest descent method to deconvolve reflectance spectra of silicates with a number of Gaussians and a continuum as a function of wavelength. Their approach, MGM, has been shown capable of deconvolving the spectra of olivine [e.g., 3] and Mg-Fe pyroxenes [e.g., 4,5] very successfully and obtains clear correlations between the central wavelengths of deconvolved Gaussian bands and chemical composition of minerals.

Although the steepest descent method is a very powerful mathematical method to find a minimum of a simple function but is not necessarily good at finding the global minimum of a complex function with many local minima. Depending on the initial choice of parameters, an obtained result may be just a local minimum of E. Furthermore, this mathematical procedure needs to know how many model parameters (e.g., the number of Gaussians for MGM) should be used before the analysis. Generally, a fitting with more Gaussians will lead to a smaller E, but the complexity of the synthetics may become too large (e.g., over fitting). Our recent analysis indicates that the use of different numbers of Gaussians for MGM analysis would dramatically change the result [6], underscoring the importance of the choice of the number of Gaussians.

Because of these problems, one must examine one by one if an MGM code properly fits an observed spectrum, ensuring the solution is not trapped in a local minimum or over fitting the data with excessive number of Gaussians. Thus, automatic spectral fitting with MGM algorithms has been very challenging.

New MGM Algorithm: In order to resolve these problems, we derived a new MGM algorithm using a Bayesian estimation approach, the exchange Monte Carlo (EMC) method, and the annealing method in this study. Because the mathematical details and numerical validation of the algorithm used in this study are given by [7,8], we discuss very briefly the concept of the Bayesian and Monte Carlo methods here.

Bayesian estimation. When the true spectrum is a combination of K of Gaussians, the most likely K value maximizes the posterior probability $P(K|Y)$ of occurrence of K for a given dataset Y:

$$P(K|Y) \propto \int P(K) \exp\left\{ -\frac{n}{\sigma^2} E(\mu_k, K) \right\} P(\mu_k) d\mu_k \quad (3)$$

where n and σ are the number and the standard deviation of data, respectively [e.g., 9]. $P(K)$ and $P(\mu_k)$ are the prior distributions of K and μ_k, respectively, and are constant or very weak functions within their possible parameter ranges. Although the expression may appear complicated, the integral in eq. (3) is numerically straightforward. The K value giving the highest value of the integral is the optimum number K of Gaussians.

Monte Carlo method: In order to resolve the problem of local minima, a probabilistic algorithm is used; the deterministic eq. (1) is replaced by eq. (4) involving random number *rand*(m) in the Monte Carlo method:

$$\mu_k^* = \mu_k^m - rand(m)\varepsilon . \quad (4)$$

This probabilistic iteration allows a solution trapped in a local minimum to escape there and reach the global

minimum. Combining with the annealing method, this probabilistic approach dramatically reduces the dependence of the initial parameter values, such as band central wavelengths, in MGM analyses.

Validation Analyses: In order to examine the validity and applicability of the new MGM algorithm proposed in this study, we conducted actual spectral deconvolution analyses using a series of reflectance spectra, olivine powder samples with different Mg/(Mg+Fe) ratios (mg#) and olivine-pyroxene mixtures with different mixing ratios. The reflectance spectra were obtained from NASA's RELAB website. Here, note that the olivine reflectance spectra used in this study are different from those used by [4].

Olivine Results: The results of our Bayesian MGM analysis indicate that the optimum number of Gaussian bands to reproduce the complex absorption band around 1 μm predicted is three for the reflectance spectra of all the olivine samples (Fig. 1). This optimum number is the same as the number estimated empirically by [4]. This coincidence is important because the number of Gaussians has been very difficult to determine objectively but highly influential to the model parameters to characterize individual Gaussian bands. This agreement between this study and previous fitting results by [4] strongly supports the validity of our new MGM algorithm

Furthermore, the central wavelengths of individual bands of the 1-μm olivine band increases gradually as mg# decreases; more iron leads to longer wavelengths (Fig. 2). The band widths and the relative intensities of the three individual bands are relatively constant regardless of mg#. All these parameters of individual Gaussian bands obtained in this study are practically the same as those obtained by a conventional MGM [4]. This supports that the results of these MGM analyses are not specific to particular datasets or analytical methods but reflect the intrinsic optical properties of olivine.

Pyroxene-Olivine Mixture Results: Though not all Gaussians are detected, most of the deconvolved Gaussians are close to one of the Gaussians of the end-member samples (Fig. 3). This will allow us to identify the mineral components within such mixtures. In particular, the bands around 1.2 μm and 1.03 μm in the mixtures of olivine can be detected when its mixing ratio is as low as 25%.

Conclusion: The above results along with the intrinsic properties of the new algorithm (e.g., little dependence on initial parameter value selection and the capability of finding the optimum number of Gaussians) suggest that method would be appropriate for automated analyses and greatly expand the applicability of MGM greatly, particularly for large volume of spectral datasets obtained for Moon.

References: [1] Pieters, C.M. and P.A.J. Englert (1993) *Remote geochemical analysis: Elemental and mineralogical composition*. [2] Sunshine, J.M. et al. (1990) *JGR*, *95*, 6955. [3] Sunshine, J.M. and C.M. Pieters (1993) *JGR*, *98*, 9075. [4] Sunshine, J.M. and C.M. Pieters (1998) *JGR*, *103*, 13,675; [5] Klima, R.L. et al. (2007) *MAPS*, *42*, 235. [6] Tsuboi, N. et al. (2009) *Proc. 42th Lunar Planet. Symp*., 118. [7] Nagata, K. et al. (2011) *Neural Networks*, to be submitted. [8] Sugita, S. et al. (2011) *JGR,* to be submitted. [9] Bishop, C.M. (2006) *Pattern recognition and machine learning*.

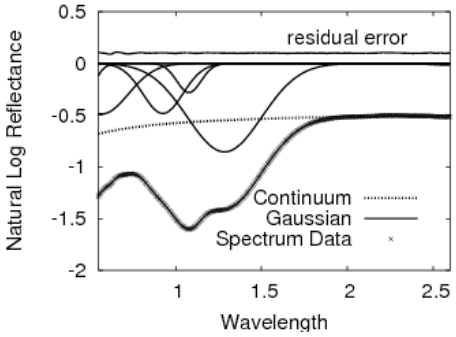

Fig. 1. The deconvolution result of an olivine spectrum using the new MGM algorithm.

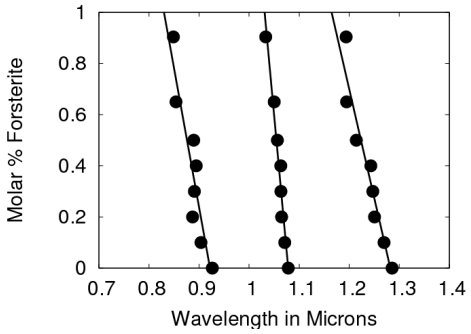

Fig. 2. The trend of central wavelength of deconvolved absorption bands of olivine with a variety of mg#.

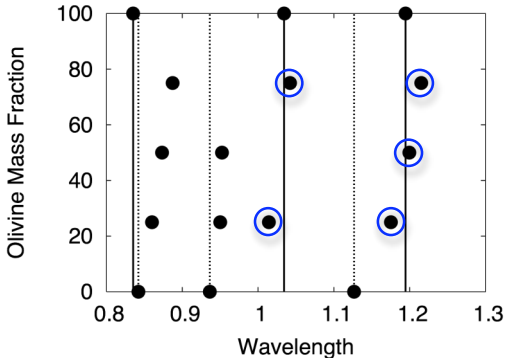

Fig. 3. The central wavelengths of Gaussians of olivine-pyroxene mixtures deconvolved with the new MGM algorithm. The solid and dashed vertical lines indicate both olivine and pyroxene band centers, respectively. As shown with blue circles, olivine absorption bands are correctly detected down to 25wt% of mixing ratio.