DERIVING LUNAR MINERAL ABUNDANCE WITH HYPERSPECTRAL REFLECTANCE DATA. L. Li[1] and S. Li[1], [1]Department of Earth Sciences, 723 W. Michigan Street, Indiana University-Purdue University, Indianapolis, IN, 46202; emails: ll3@iupui.edu and shuali@imail.iu.edu.

Introduction: Information on the spatial distribution of lunar surface minerals has a number of applications in investigating lunar surface processes and geological evolution and in planning future missions. Previous efforts of mapping lunar mineral abundance have been focused on the use of Clementine ultraviolet (UV), visible (VIS) and near infrared (NIR) multispectral images. With the release of the M3 images [1], it is expected that the algorithm for mapping lunar mineralogy should be required. Although some algorithms for such mapping purpose exist [2-4], they are either time consuming or difficult to apply to hyperspectral images. Here we introduce an approach to mapping major lunar minerals with hyperspectral reflectance, in which genetic algorithms (GA) and partial least squares (PLS) regression are jointly used. GA-PLS has been trained with the Lunar Soil Characterization Consortium (LSCC) dataset [5] and can be directly applied to the M3 reflectance images for mapping major lunar minerals.

Methods: We used the LSCC spectral dataset (300-2600 nm, 5 nm spectral resolution) resampled into the M3 spectral resolution (430-3000, 10 nm spectral resolution) [1] to build the spectral-mineral models, in which 57 LSCC samples from particle size groups: 45-20, 20-10, <10 μm were used to train GA-PLS models. These models were then validated on the spectra of the <45 μm particle size group. Because the mineral abundances for this bulk soil group are not available, the abundance values of a mineral were averaged across samples from particle size groups: 45-20, 20-10, <10 μm and used to approximate the abundance of the same mineral for the corresponding bulk soil sample.

GA-PLS uses PLS to create training models for deriving mineral abundance from reflectance spectra and appeals to GA to select a subset of spectral bands that are sensitive to a mineral compositional parameter. Briefly, a simple PLS model consists of two outer relations and one inner relation. The two outer relations result from eigenstructure decompositions of both the matrix containing explanatory variables (i. e., spectral bands) and the matrix containing response variables (i. e., lunar mineral abundance), while the inner relation links the resultant score matrices from the two eigenstructure decompositions generating the outer relations [6]. Let both X [$n\times m$] represent a explanatory matrix, the first outer relation is derived by applying principal component analysis (PCA) to X, resulting in the score matrix T [$n\times a$] and the loading matrix P' [$a\times m$] plus an error matrix E [$n\times m$], i.e. X = TP' + E Similarly, the second outer relation for Y [$n\times p$] standing for a response variable matrix can be derived by decomposing Y into the score matrix U [$n\times a$] and the loading matrix Q' [$a\times p$] and the error term F [$n\times p$], i.e. Y = UQ' + F. The prime represents matrix transpose. The inner relation U = BT is a multiple linear regression between the score matrices U and T in which B is an $n\times n$ regression coefficient matrix determined via least square minimization. The goal of the PLS model is to minimize the norm of F while maximizing the covariance between X and Y by the inner relation.

A genetic algorithm includes at least five components: encoding, population initialization, individual selection, crossover and mutation [7]. For encoding a genetic algorithm works with a population of randomly generated chromosomes (i.e. candidate solution) in population initialization, and each chromosome is formed by as many "bits" as the number of spectral bands, and "zeros' or "ones" are assigned to the bits of that chromosome. A zero bit indicates the band is not selected and an one bit means otherwise. Crossover is the process of reproducing new offspring by exchanging genes between two reproducing chromosomes, and then mutation follows and simulates the gene change of a chromosome due to the random disturbance. Because of its capability of removing spectral bands that are insensitive to a response variable, GA enables PLS to build a simple training model from hyperspectral dataset, for which only a few spectral bands are used [8].

Results and Discussion: The result from GA-PLS modeling of major lunar mineral abundances is summarized in table 1 and shown in figure 1 are correlations between estimated and measured abundances only for agglutinate, pyroxene and olivine. For training models, GA-PLS resulted in a high coefficient of determination for agglutinate (R^2= 0.78), pyroxene (R^2= 0.89), plagioclase (R^2= 0.91), ilmenite (R^2= 0.95), olivine (R^2= 0.58); for model validation, the coefficient of determination is 0.73 for agglutinate, 0.92 for pyroxene, 0.93 for plagioclase, 0.96 for ilmenite and 0.70 for olivine. A relatively poor correlation for olivine is similar to that we reported last year [9] and may stem from a narrow range of its abundances and the fact that the LSCC samples are not representative.

Table 1. R^2 values for training and validating GA-PLS

Minerals	Training	Validating
agglutinate	0.78	0.73
pyroxene	0.89	0.92
plagioclase	0.91	0.93
ilmenite	0.95	0.96
olivine	0.58	0.70

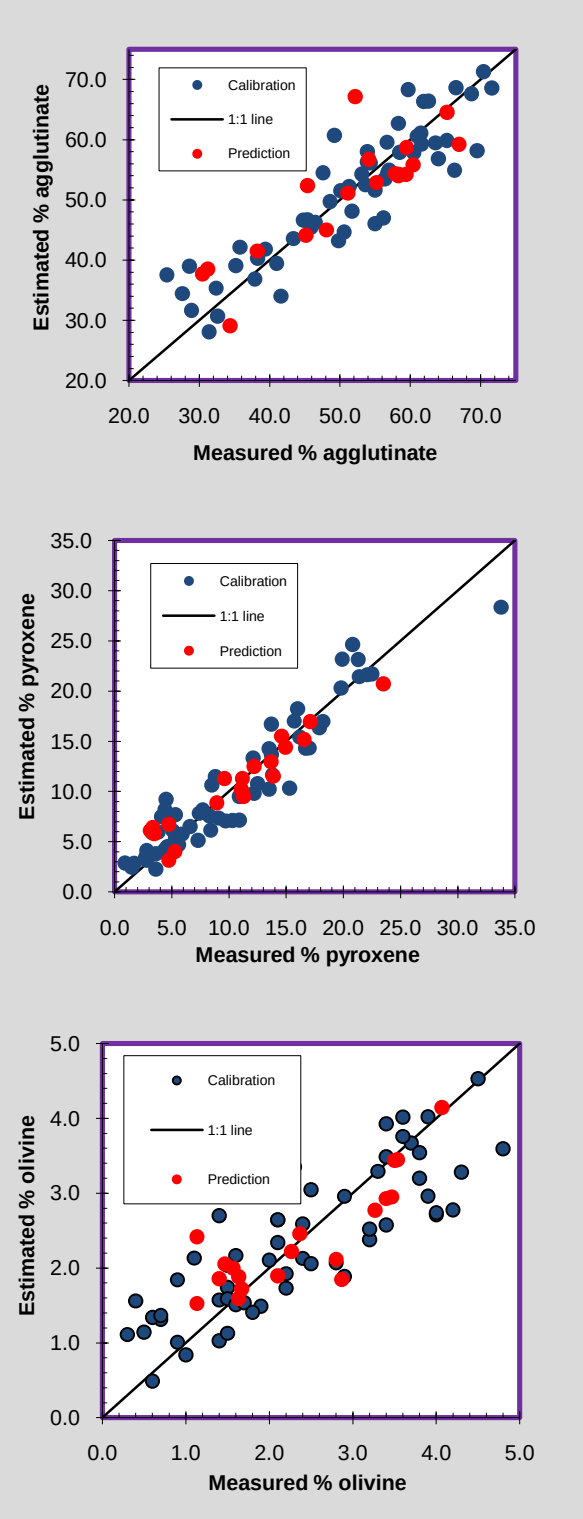

Figure 1. Comparison between estimated and measured abundances for agglutinate, pyroxene and olivine from top to bottom.

Observing high correlations between estimated and measured mineral abundances shown in figure 1, one may be curious what the GA-PLS model looks like for each of major lunar minerals. While GA-PLS is often built upon the inner relationship between latent variables for both spectral data and compositional parameters. For the purpose of practical mapping applications with the M3 hyperspectral data, the derived GA-PLS models have been converted into regressions of spectral reflectance at different bands to the measured abundance of a mineral. The regression relations for agglutinate, pyroxene and olivine are shown in table 2.

Table 2. Regression relationships for estimating lunar minerals

Minerals	Regression equation*
agglutinate	$C = 634.3R_{439}-364.1R_{449}-90.4R_{459}-164.8R_{469}+446.3R_{2238}-1354R_{2298}-1272.9R_{2308}+2166.5R_{2318.2}+8.55$
pyroxene	$C = 234R_{683}+164.4R_{693}+40.9R_{702}-144.8R_{712}-322R_{721}+239.4R_{2258}-205.3R_{2437}-4.498$
olivine	$C = -0.721R_{438}-146.89R_{448}+148.65R_{458}+2.892$

*: R_{xxxx} is log (1/reflectance) at a specified wavelength (xxxx nm) and C is estimated concentration for a mineral.

Regressions for plagioclase and ilmenite were also derived, but not shown here because more than 30 spectral bands were used for estimating these minerals. This is due to the fact that both minerals don't have strong diagnostic absorption features and many bands could be selected by GA as PLS regression variables.

The regression relationships shown in table 2 are for the decimal reflectance spectra with the natural logarithmic transformation, the M3 images have to be calibrated into decimal reflectance and logarithmically transformed before these models could be applied for mapping lunar minerals.

Conclusions: Application of GA-PLS generated spectral-compositional models for estimating major lunar minerals from hyperspectral reflectance data. These models can be directly applied to the M3 hyperspectral images for lunar mineralogy mapping after the M3 images are calibrated to decimal reflectance and logarithmically transformed. While sophisticated methods for the M3 data preprocessing are available, implementation of these methods are not as straightforward as natural logarithmic transformation and the latter should thus be preferred for a practical mapping purpose with hyperspectral datasets. Further efforts are to adapt these models for mapping lunar mineralogy with the hyperspectral M3 images.

References: [1] Pieters C., et al. (2007) *Lunar and Planetary Science* XXXVIII, # 1295. [2] Sunshine J. M. and Pieters C. M. (1993) *JGR*, 98**,** 9075–9087. [3] Lucey P. G. (2004) *Geophys. Res. Lett.*, doi:10.1029/2003GL019406. [4] Lawrence S. J. and Lucey P. G. (2007) *JGR*, doi:10.1029/2006JE002765. [5] Taylor L.et al. (2001) *J. Geophys. Res.* 106, 27,985–28,000. [6] Geladi P. and Kowalski B. (1986) *Analy. Chimi. Acta*, 185, 1-17. [7] Forrest S. (1993) *Science*, 261, 872-878. [8] Li et al. (2007) *IEEE Trans. Geosci. and Remote Sens. Lett.*, 4(2), 216-220. [9] Li L. and Li S. (2010) *LPSC XXXXI*, abstract # 2189.