

QUANTIFYING TiO₂ ABUNDANCE OF MARE SOILS: A STRATIFIED PARTIAL LEAST SQUARES APPROACH TO QUALITATIVE INTERPRETATION. L. Li.¹ and P. G. Lucey², ¹Indiana University-Purdue University, 723 W. Michigan Street, Indianapolis, IN 46202, email:ll3@iupui.edu, ²University of Hawaii, Institute for Geophysics and Planetology, 1680 East-West Road, POST 515, Honolulu Hawaii, 96822.

Introduction: Improved remote estimation of lunar soil TiO₂ abundance has important implications for characterizing lunar mare compositional diversity as well as addressing basaltic volcanism [e.g., 1-5]. Quantifying lunar soil TiO₂ abundance with visible to near infrared spectral data is commonly achieved by relating spectral characteristics of lunar soil samples returned by the Apollo and Luna missions to their TiO₂ abundance. Traditional methods utilize spectral bands and/or band ratios for this estimation [e.g., 2, 4, 6-12], while two multivariate data analysis methods, principal component regression (PCR) [13-15] and partial least squares (PLS) regression [15-16] have been recently used. A difficulty with applying PCR/PLS is to explain the spectral-compositional relationships, i.e. how TiO₂-bearing minerals governs the models. Spectroscopic analyses of the PLS and PCR eigenfunctions (eigenvectors) used in previous studies [13-14] may not be appropriate because PLS differs from PCR in the way to derive its spectral eigenfunctions. As identification of the mineral causal effect has important implications for improving the estimation of lunar soil TiO₂ abundance, especially for low TiO₂ (<4%) samples [17], here we report the use of PLS regression to estimate TiO₂ abundance and provide an insightful interpretation of PLS results by applying stepwise multiple regression analysis (SMRA) to the PLS components.

Partial Least Squares (PLS) Modeling: A simple PLS model consists of two outer relations and an 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., soil chemical abundance), while the inner relation links the resultant score matrices resulting from the two eigenstructure decompositions [16, 18]. The inner relation is often represented by a multiple linear regression between the score matrices and the corresponding regression coefficients are determined via iterative least squares minimization [18]. Like PCR used in [13-15], PLS assumes that the spectra-composition system of interest is driven by a few of dominant factors or components and that their corresponding scores can be used to predict a response variable; Unlike PCR, however, PLS has the advantage of determining a small set of dominant factors that not only explains the variance of the explanatory variable but also has high correlation to the response variable. The approach to the PLS solution can be found in [16, 18].

Application of PLS to LSCC Dataset: The Lunar Soil Characterization Consortium (LSCC) group generated four sample replicates for each of the 9 mare soil samples at the particle size separate: <45, 45-20, 20-10, and <10 μm . TiO₂ and major mineral (agglutinate, pyroxene, plagioclase, olivine, ilmenite, and volcanic glass) abundance of these replicates were determined. Besides these major minerals, each subsamples also includes other constituents such as unclassified and FeS minerals which are represented by the remaining component (RC) here. The reflectance spectra of these replicates measured by the RELAB of Brown University [15].

For PLS modeling the LSCC mare samples were used and stratified into two subsets: high-Ti mare soils from the Apollo 11 and Apollo 17 sites, and low-Ti mare soils from the Apollo 12 and Apollo 15 sites. Each subset was split into the training set for training PLS models and the validation set for model assessment, i.e. that the TiO₂ abundance and reflectance for three particle size groups: 45-20, 20-10, and < 10 μm was used as the training set, while the data for bulk soil, i. e. <45 μm in particle size were used for validation. The LSCC reflectance spectra in the spectral range 0.4-2.5 μm were resampled into a coarse spectral resolution (nominal 10 nm intervals). The resampled spectra were then further converted to absorbance spectra by the natural logarithmic operation of each spectrum.

Results: The root mean squares error (RMSE), coefficient of determination (R^2) resulting from PLS modeling of the LSCC mare spectra for estimating TiO₂ abundance are presented in the following table along with the number (#) of PLS factors (PLSF):

Samples	RMSE [*]	R ²	PLSF
Ap12 & Ap15 (low-Ti)	0.043/0.21/8.3%	0.99	8
Ap11 & Ap17 (hi-Ti)	0.21/0.50/6.3%	0.97	7
Low-and hi-Ti samples	0.1/0.77/14.04%	0.94	14

*: the first value is rmse for calibration, and the second and third values are rme and relative error for validation; R^2 is for calibration.

Correlation between the PLS estimated and measured TiO₂ abundance is shown in Figure 1 for the low-Ti mare samples and Figure 2 for the high-Ti samples.

Discussion: The PLS result (see Table) shows that estimated TiO₂ has a larger relative error (8.3%) for validation on the low-Ti mare samples than that (6.3%) for the high-Ti mare samples, but requires more PLS factors for the low-Ti dataset than those for the high-Ti dataset. Application of SMRA to the regression of

TiO₂ abundance against the abundance of seven major minerals indicates that the dominant TiO₂-bearing minerals are agglutinate, pyroxene and ilmenite for the low-Ti samples, but ilmenite and RC are the most important TiO₂-bearing minerals for the high-Ti samples. This result is consistent with the interpretation in [1]. However, these SMRA results are not in agreement with the PLS regression results; otherwise, three and two PLS factors rather than eight and seven should have been used for estimating TiO₂ of low- and high-Ti mare samples respectively.

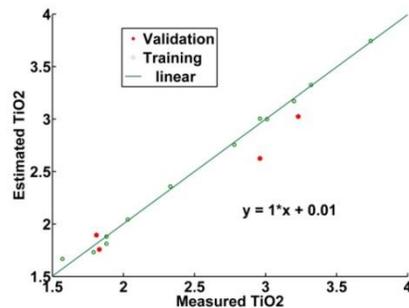


Figure 1. Correlation between estimated and measured TiO₂ for low-Ti mare samples.

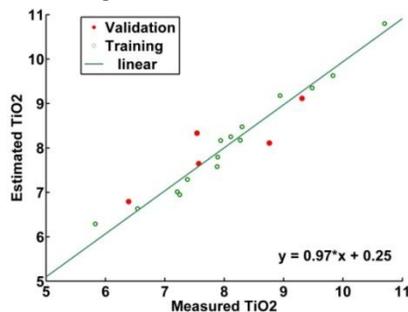


Figure 2. Correlation between estimated and measured TiO₂ for high-Ti mare samples.

To address this inconsistency between the number of PLS factors and the number of dominant TiO₂-bearing minerals, SMRA were used in the regression of the spectral score for infidel PLS factors against the abundance of TiO₂-bearing minerals. For low-Ti samples, the SMRA analysis indicates that the 1st PLS factor is driven by and positively correlates to agglutinate, pyroxene, plagioclase, ilmenite and RC and the 2nd is related to pyroxene. However, the 3rd to 5th and 8th factors don't show significant correlation, and they represent the effect of physical parameters (e.g. particle size). The 6th and 7th factors show negative correlations to plagioclase and volcanic glass respectively; these two are explained to compensate the effect of plagioclase on the 1st factor. This explains why low-Ti samples have 3 dominant Ti-bearing minerals but 8 PLS factors are required to achieve good Ti estimation.

Similarly, the SMRA for the high-Ti samples suggests that the 1st PLS factor is driven by and positively correlates to agglutinate, pyroxene and ilmenite; the 2nd PLS is positively related to agglutinate, pyroxene, volcanic glass, plagioclase, ilmenite and RC; and the 3rd PLS factor is positively related to plagioclase and ilmenite. It is evident that for the high-Ti soils suggests that the spectral information of the dominant TiO₂-bearing ilmenite was distributed across the first three PLS factors but the spectral signal for RC was concentrated in the 2nd PLS factor. For the high-Ti samples, SMRA suggests that TiO₂ abundance is driven by ilmenite and RC. This implies that to compensate the effect of agglutinate, pyroxene, volcanic glass, plagioclase on the first two PLS factors and of plagioclase on the 3rd factor, additional PLS factors are required. This is why the fourth PLS factor was included in the PLS model, which is negatively related to agglutinate, pyroxene, volcanic glass and plagioclase. However, the 5th and 7th factors don't show significant correlation to mineral abundance and their inclusion be due to the effect of physical parameters (e.g. particle size). This explains why the PLS model for high-Ti samples needs 7 PLS factors other than 2. Unlike the case for low-Ti samples, the effect of physical property is present in the low rank factors for high-Ti soil samples, implying the effect is not as important as TiO₂-bearing ilmenite.

Conclusion: PLS regression was applied to the LSCC low- and high-Ti mare samples and resulted in high correlation between the estimated and measured TiO₂. SMRA was used for qualitative interpretation of the PLS factors and demonstrated as an efficient approach to the interpretation of the PLS spectral-compositional relationships.

Acknowledgement: This work is partially supported by NASA LASER grant NNX08AZ06G (PGL).

References: [1] Pieters C. M. (1978) PLPSC, 1978-2849. [2] Pieters C. M. et al. (1993) JGR, 98,17,127-17,148. [3] Giguere T. et al. (2000) Meteoritics & Planet. Sci., 35, 193-200. [4] Gillis J. J. et al. (2003), JGR, 108, doi: 10.1029/2001JE001515. [5] Kramer G. Y. et al. (2008), JGR, 113, doi:10.1029/2006JE002860. [6] Charette M. T. et al. (1974), JGR,79, 1605-1613. [7] McCord, T. B. et al. (1981) JGR, 86, 10,883-10,892. [8] Johnson T. V. et al. (1977), PLSC 8, 1029-1036. [9] Pieters C. M. et al. (1976) PLSC, 2677-2690. [10] Blewett D. T. et al. (1997), JGR, 102, 16,319-16325. [11] Lucey P. G. et al. (2000), JGR, 105, 20,297-20,305. [12] Shkuratov G. et al. (2003), JGR, 108, doi:10.1029/2002JE001971. [13] Jaumann R. (1991), JGR, 96, 22,793-22,807. [14] Pieters C. M. et al. (2002), Icarus, 155, 285-298. [15] Li L. (2006), JGR, 111, E04002, doi:10.1029/2005JE002598. [16] Li L. (2008), JGR, 113, doi:10.1029/2008JE003213. [17] Gillis-Davis J. J. et al. (2006), Geochimica. Cosm. Acta, 70,6079-6102. [18] Geladi P. & Kowalski B. (1986) Anal. Chim. Acta, 185, 1-17.