INVESTIGATING MARS' ELEMENT DISTRIBUTION WITH CRISM SUMMARY PRODUCTS

O.M. Kamps[1], R.D. Hewson[1], F.J.A. van Ruitenbeek[1], F.D. van der Meer[1], [1]University of Twente, ITC (Faculty of Geo-Information Science and Earth Observation) (o.m.kamps@utwente.nl)

Introduction: The low spatial resolution of element maps based on gamma-ray spectroscopy (GRS) [1], makes it difficult to conclude on a process behind their distribution. For the different elements, several studies have proposed theories that could explain their distribution [1,2,3], but so far for most elements, there is not an agreement on one theory yet.

This study aims to obtain a better understanding of factors that influence the element distribution by comparing GRS-element maps using data measured by the CRISM instrument in survey mode [4]. Infrared spectroscopy data from CRISM estimates the chemistry and mineralogy of the surface, which might relate to the GRS-element concentration.

This comparison is performed by using Partial Least Square Regression (PLSR) [6], using CRISM summary products [4] as the predictors and the GRS-elements - Fe, Th, K, Cl, H_2O, Si - as the response variables. It is hypothesized that such a model works best for iron and water, because many of the summary products describe absorption features that relate to these elements [4,5].

Spectral interpreted mineralogy maps have been compared before with GRS-element maps [7,8]. The difference of these studies with what is done here is that the infrared data is used to understand the element distribution instead of creating new element maps.

Method: Differences between gamma-ray and near-infrared spectroscopy must be taken into account before their data can be compared.

Firstly, the difference in penetration depth of gamma-ray and infrared rays is significant. The gamma-ray spectrometer measures the first few decimeters, while CRISM only measures the first micrometers. Therefore, when both datasets are compared, it has to be assumed that the upper layer (few decimeters) is homogeneous. This assumption does not apply when the surface is covered by dust. The dust coverage index (DCI) [9] is averaged for each GRS-pixel, where the threshold of a DCI > 0.96 [9] has been used to select the pixels that are assumed to be uncovered by dust.

Secondly, there is a major difference in spatial resolution of both datasets, 200-500km for GRS [1] and ~200m for the 5*5° tiles of CRISM images measured in survey mode [4]. To tackle the difference in resolution, the statistics - such as the mean, median, standard deviation, number of pixels – of the CRISM summary products were calculated for each selected GRS-pixel.

Processing of the CRISM-summary products is required before these can be used as predictors for the PLSR-analysis. First of all, only surface related summary products have been used for this analysis. Furthermore, the pre-processing includes removal of outliers, using normalization and standardization techniques.

To test the accuracy of the model, cross validation is applied. Two-thirds of the data was randomly selected and used for model calibration, the other third was used as validation dataset. The averages and standard deviations of the PLSR results were calculated with a monte-carlo simulation of ten-thousand repetitions.

When such a PLSR-model is created, defining the right number of predictor components is important. For example, too few components and the model is not describing enough variance, too much and the model is overfitting the data. The number of component based on the cumulative explained variance for each component.

Results and Discussion:

Coefficients of determination

For each element, the coefficient of determination (R^2) has been calculated, to indicate the variance between the modelled values and original data. Both the calibration and validation data are summarized in Table 1.

	Fe	K	Th	Cl	Si	H_2O
Number of components	5	7	4	10	7	12
Calibration						
R^2	61%	56%	44%	32%	39%	41%
R^2-error	0.6%	0.3%	0.6%	0.2%	0.3%	0%
Validation						
R^2	58%	51%	41%	23%	33%	33%
R^2-error	4%	5%	4%	4%	4%	4%

Table 1: Comparison of coefficients of determination

In Table 1, it can be noticed that iron has the highest R^2-value, followed by potassium and thorium. However, chlorine, silica, and water, have low R^2-values for both calibration and validation datasets. Especially for water this was not expected, since there are many CRISM-summary products that describe water features in the near-infrared [5]. Differences can be observed in standard deviations between the calibration and validation groups, with higher accuracy in the validation dataset.

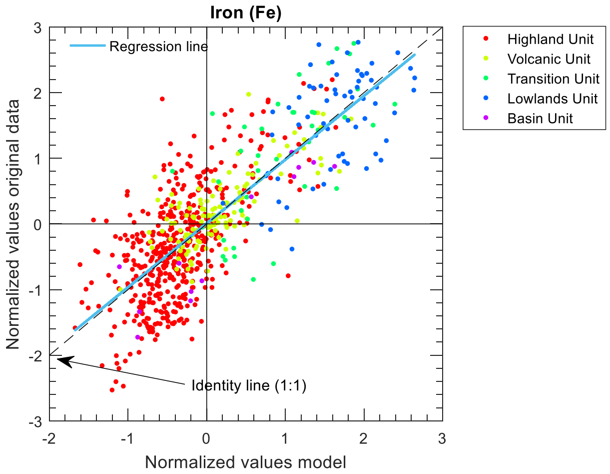

Figure 1: Scatter plot of the modelled values versus the original data, grouped per lithology

In figure 1 the modelled Fe concentration is compared with original concentration. It can be observed that the regression line plots close to the identity line. Figure 1 shows that the model works fine to distinguish the different geological units [10], with higher concentrations of iron in the lowland area's and lower concentrations in the highlands and volcanic units.

Regression coefficients

To test how different summary products contribute to the final model, the regression coefficients for each product were compared. As an example, figure 2 shows the regression coefficients for the PLSR-model for iron. As can be seen, the summary products have been grouped. This is based on hierarchical cluster analysis, after calculating the correlation coefficients.

In figure 2 it can be noticed that mainly the summary products that relate to mafic minerals (group 1 in Figure 2), have high regression coefficients, which means these contribute most to the final model. The regression coeffients of summary-products related to hydrated minerals are relatively low. The same applies for the dust related summary products. Current work focuses on the interpretation of the regression coefficients for each element. Understanding the relationship between the high regression coefficients and the elements could explain the spatial variance of the element concentration.

References: [1] Boynton, W.V. et al. (2007) *JGR, 112,* E12S99. [2] Taylor, G.J. et al. (2006) *JGR*, *111*, E03S10, [3] Keller, J.M. et al. (2006) *JGR*, *111*, E03S08. [4] Pelkey, S.M. et al. (2007) *JGR*, *112*, E08S14 [5] Viviano-Beck, C.E. et al. (2014) *JGR: Planets*, *119*, 1403-1431 [6] Wold, S. et al. (2001) *Chemometrics and Intelligent Laboratory System*, 58, 109-130 [7] McSween Jr., H.Y. et al. (2009) *Science*, *324*, 736-739 [8] Riu, L. et al. (2017), *LPS XLVIII*, Abstract #1937. [9] Ruff, S.W. et al. (2002), *JGR*, *107*, E12. [10] Tanaka, K.L. et al (2014), USGS *SIM3292*

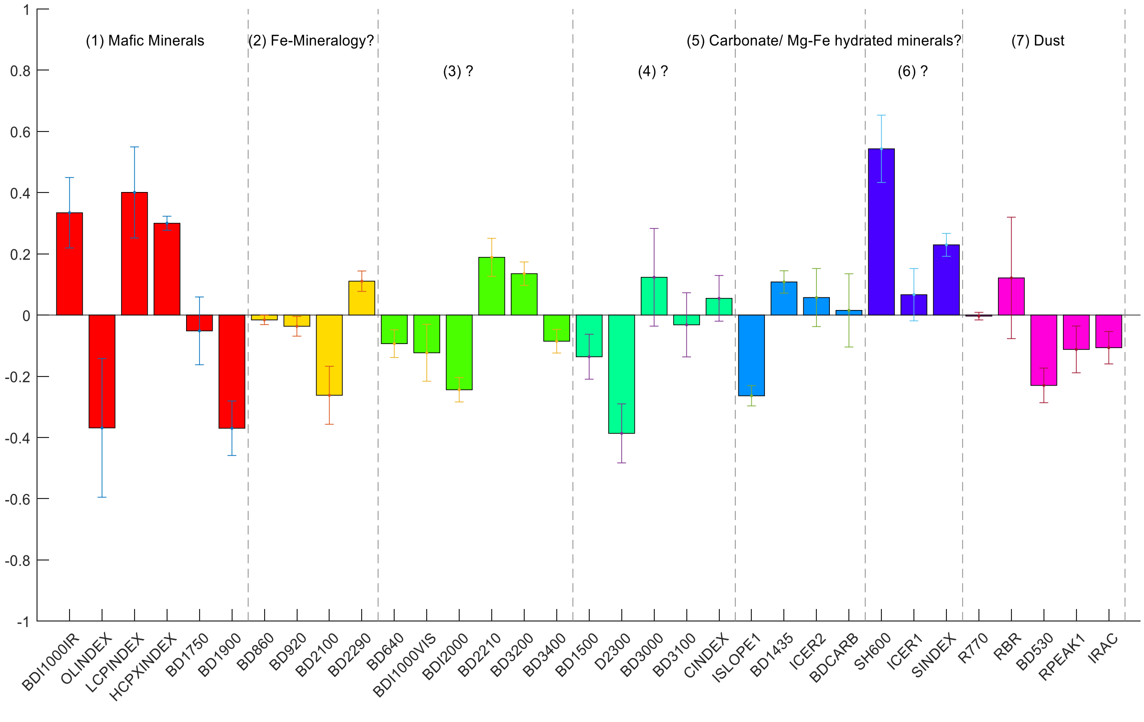

Figure 2: Beta-values of the linear regression model of iron. Bars present the mean, error bars the standard deviation of the variation after thousand repetitions