© EARSeL and Warsaw University, Warsaw 2005. *Proceedings of 4th EARSeL Workshop on Imaging Spectroscopy. New quality in environmental studies*. Zagajewski B., Sobczak M., Wrzesień M., (eds)

USING IMAGING SPECTROSCOPY FOR THE QUANTITATIVE DETERMINATION OF SOIL IRON CONTENT IN PARTIALLY VEGETATED AREAS

Harm Bartholomeus, Gerrit Epema and Michael Schaepman

Centre for Geo-Information, Wageningen University and Research Centre, Wageningen, The Netherlands, Harm.Bartholomeus@wur.nl

ABSTRACT

We determine the soil iron content in a semiarid region of Spain using airborne imaging, field and laboratory based spectroscopy. By assessing the influence of vegetation fractional and fully covered areas are separated for specific analysis each. The derived distribution of soil iron content is spatially discontinuous so spatially discontinuous quality indicators are assigned depending on the retrieval accuracy of iron content.

The test site is located at the slopes of the El Hacho Mountain, near Álora in Southern Spain, showing a large variety in iron content. Orchards with olive trees are the dominating land use of this site. There are significant gaps between vegetation covered and bare soil patches, which are assessed using a spatial resolution that allows mapping of three categories, bare soil, partly vegetated and fully vegetated areas.

Since the accuracy of the iron prediction depends on the gap fraction or fractional vegetation cover (fCover) of the orchard, a spectral unmixing based approach is used to discriminate several classes of vegetation cover. The iron content of areas with a low fractional vegetation cover can be determined accurately using a quantitative statistical based model. If the vegetation fraction is significant, accurate determination of the iron content is not possible, and consequently these areas need to be interpolated or excluded from the spatial iron map. For the areas with an intermediate vegetation cover, correction of the mixed spectrum is performed, making use of the vegetation fractions determined in the first step and a vegetation spectrum derived from olive trees. This results in a residual soil spectrum, which is used to determine the soil iron content with intermediate quality.

The final result is a spatial distribution of iron content with prediction uncertainty. The user of the geographical database can determine which points may be used during further analysis, depending on the required accuracy.

INTRODUCTION

The slopes of the El Hacho mountain complex near Álora – Southern Spain show a large variety in soil iron content. However the role of iron in plant physiology is not fully understood iron is important for agriculture. It is involved in chlorophyll synthesis and is a component of other plant tissues [1]. Both excessive and deficient iron content of soils, and the form in which it occurs, can make agricultural crops vulnerable to both qualitative and quantitative damages. Furthermore the iron concentration, and directly connected to this the soil colour, is an indicator for the age of the soil, which has a large influence on the fertility.

Determining the spatial distribution of different types of iron with traditional fieldwork and laboratory analysis is time-consuming and expensive [2]. Remote Sensing has proven to be a useful tool for determining the presence of iron in large areas and various research fields [3]; [4]; [5]. Large areas can be mapped and soils are left undisturbed during sampling.

The influence of iron on the electromagnetic spectrum has been widely investigated and it is shown that quantification of the amount of iron with spectral measurements is possible e.g. [6]; [7]; [8]. However, research is often limited to practically bare soil areas. Vegetation has a large influ-

ence on the soils reflectance spectrum, and results in highly inaccurate iron quantity estimations [9].

El Hacho Mountain consists of Tertiary Marls and sands, deposited on a continental slope, which are indicated with the generic term flysch deposits. On top thick cemented deposits of sand and gravel (conglomerate) occur [10]. Conglomerate blocks are now scattered over the flysch slope, where Cambisols are formed. Behind the blocks unweathered flysch is exposed. Iron content is varying with the distance down slope from these blocks, due to variation in erosion, iron deposition and leaching. The presence of gullies further increases the spatial heterogeneity. Hematite is the major iron mineral in the area, while just behind the conglomerate blocks some goethite may be present.

METHODOLOGY AND MATERIALS

Soil Samples

In summer 2003, 35 bare soil plots are sampled and measured for iron content. The samples are positioned in two down slope transects and cover the variation over the slopes. Fractional cover of soil, rocks and vegetation is determined and a mixed sample of the topsoil (0-2 cm) collected. In the major study area for all plots, four field spectra and samples for determination of iron content and measuring laboratory spectra are taken. The plots are divided in a training set (19 plots) and a reference set (16 plots), to allow cross-validation of the results.

The total iron concentration of all 35 soil samples is determined using a dithionite extraction [11]. The concentration of iron in the extraction fluids is determined with an ICP-AES.

Field and laboratory spectra

Field and laboratory spectra are acquired with an ASD Fieldspec Pro FR, covering the 350 – 2500 nm wavelength region. For the field spectra four measurements per plot are taken of the surface. The field of view was 25° and measurements are taken from nadir at a distance of about 30 cm from the surface. The soil samples used for the laboratory analysis are air-dried and sieved at 2 mm. The incidence angle of the lamp is set to 30° of nadir and a 3° fore optic is used at a 30 cm distance from the target. The sample is rotated 90° in between the four measurements, which are done per sample. For further analysis these reflectance field and laboratory spectra are limited to and rescaled to the bandwidth of ROSIS.

Image data

Airborne hyperspectral data are collected in June 2001, during the DAISEX campaign [12]. ROSIS is build for the detection of fine spectral structures especially in coastal and inland waters [13], but it is also used for land applications[14]. The processing of ROSIS images to reflectance values is done by DLR, using the ATCOR model[15].

Due to the coarse resolution of the used DEM for ortho rectification, the image geometry within the test area had high distortions. The use of a handheld GPS for determination of the location appeared to be inaccurate too. Therefore the positioning of the plots on the ROSIS image is done visually. In this way correct spectral information was achieved for the sampling points. This has no influence on the spectral information and further analysis.

Iron content

Laboratory and field measurements are used to establish the regression between total iron content and reflectance in the visible and near-infrared wavelength range. The properties of the absorption dip in the visible part of the electromagnetic spectrum are used for the quantification of the iron content. Continuum removal [16] is used to normalize the reflectance, so the total absorption, described with the area of the absorption dip around 550 nm (referred to as: Area D_{550}) and the Standard Deviation of the continuum removed values in the same wavelength range can be used to estimate the iron quantity. In order to determine the most suitable technique for the final iron map-

ping the sensitivity of these methods towards fractional vegetation cover is investigated by mixing a pure soil and olive spectrum, simulating varying vegetation abundances.

Processing

Linear Spectral Unmixing [17], also named Spectral Mixture Analysis (SMA), is used to estimate the fractional covers of seven endmembers in the image. SMA is used to find the fractions (abundances) of a number of endmembers that best explain the recorded mixed pixel reflectance spectrum. SMA is a physically based model that provides quantitative results concerning the distribution of materials within an image scene [18].

The two major endmembers are the iron rich soil and olive. First the areas with an abundance of more than 10% for the iron rich soil are selected. Next the olive fractions are used to define three quality classes for the iron determination. Areas with a low fCover for olive (< 5%) are considered bare, so determination of the iron content can be done without pre-processing. For areas with fCover between 5 and 50% a residual linear unmixing model is used, which is described further on. Parts of the study area with a high olive abundance (fCover > 50%) are excluded from the analysis. The influence of vegetation is too high to determine soil parameters with a reasonable accuracy. Additional endmembers are selected to exclude areas with a high abundance of other materials.

For the areas with medium partial vegetation cover the effect of vegetation on the spectral signal has to be removed. This is done with an adapted form of the linear unmixing algorithm, which gives the residual soil spectrum:

$$R_{soil} = \frac{R_{pixel} - (fCover \cdot R_{olive})}{(1 - fCover)} \quad (1)$$

in which:
- R_{soil} = Residual reflectance of the pixel
- R_{pixel} = Actual reflectance of the original pixel
- R_{olive} = Spectral signature of the olive endmember
- $fCover$ = Olive abundance within the pixel

RESULTS AND DISCUSSION

Preprocessing

The ROSIS image shows striping in the lower bands. Moreover the spectra show a noisy pattern. Therefore MNF-transformation is used to remove the striping effect and smooth the spectra [19]. Bands 15 to112 are used for these transformations and further analysis, since the severe striping in the deleted bands could not be removed during the MNF transformation. The first seven MNF bands are used for the inverse transformation.

Linear unmixing is applied on the Forward MNF transformed images, since this algorithm can best be applied on uncorrelated datasets [20].

Iron content

For the quantification of the iron content the properties of the wide absorption dip around 550 nm (D_{550}) are used. The relation between the reflectance and the total iron content, established using laboratory measurements shows comparable R^2-values (for the Area of D_{550} and the Standard Deviation of D_{550}, both after Continuum Removal (table 1). The correlation coefficients and prediction accuracy for the training set are structurally lower than the values of the reference set. This is the result two points that are located on top of a conglomerate block where the soil has not developed deep enough to grow olives. For this reason the farmer did not remove dead organic material. This does not make a difference in the regression formula that is used for the iron prediction.

Table 1: correlation coefficients and error in iron prediction using laboratory spectra

	Correlation (R^2)		Error in Iron prediction (mass %)	
	Training	Reference	Training	Reference
Area of D_{550}	0.36	0.65	3.3	2.3
StDev of D_{550}	0.40	0.67	3.1	2.1

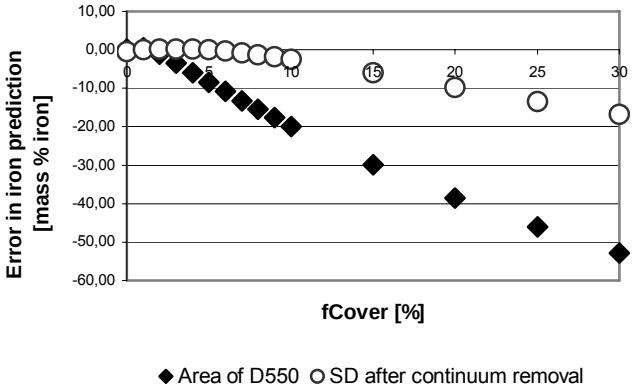

Figure 1: Relation between the simulated fCover and the error in iron prediction. The total iron content is underestimated when fCover increases.

However the iron prediction accuracy with the Standard Deviation just performs little better than the Area of the absorption dip, there is a clear reason why this regression formula is used further, which is illustrated in figure 1. This figure shows the error in iron prediction when an olive spectrum and soil spectrum are mixed simulating various fractional covers. The Area of the absorption dip shows a much higher sensitivity to the influence of vegetation. Therefore the relation based on the Standard Deviation of the Continuum Removed values of D_{550} is used for the determination of the iron quantity.

UNMIXING

The study area is unmixed using 7 endmembers. Besides the two endmembers for olive and iron rich soil, additional endmembers are selected. These consist of two additional soil types, bare rock, pavement and bushes.

First the abundance images are used to delimit the area with iron rich soil and mask the areas with a high abundance of endmember other than olive or iron rich soil.

Next the fractional olive cover is used to determine which further processing is applied on each pixel. Pixels with a fCover of less than 5 % are considered being pure soil, where no pre-processing is required before iron quantification. From the pixels with a fCover between 5% and 50% the spectral influence of vegetation is removed, using residual linear unmixing. The effect of vegetation on areas with an olive fCover above 50% is considered too big to be removed and therefore quantification of the iron content is not possible in these areas.

This results in a map with three classes that show the level of reliability (figure 5) of the accompanying soil iron content map. The user can decide which reliability is required for his analysis.

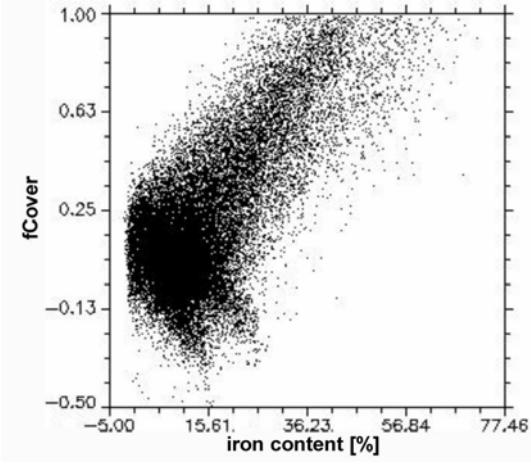

Figure 2: Scatterplot of the predicted total iron content and the fractional vegetation cover. The iron content is overestimated when fCover increases

The effect of fractional vegetation cover on the determined iron quantity is shown in figure 2. It shows the increase of the predicted iron quantity when fCover increases, resulting in an overestimation of the total iron content. The overestimation leads to iron concentrations of more than 70 %, where in the study area iron concentrations above 20 % are rare.

Residual unmixing

In the area with fCover between 5 and 50 % the effect of vegetation on the spectral response is removed, using a two-endmember linear mixture model. Inputs for the model are the MNF-denoised ROSIS image, the Olive endmember spectrum and the Olive fCover map. Figure 3 shows the spectrum of a pixel with fCover of 0.32, before and after residual unmixing. The original spectrum shows the typical effect of vegetation with the strong increase in reflectance just above 700 nm and the lower reflectance in the visible wavelengths due to absorption of the radiance by the plant tissue. After residual unmixing the influence disappeared, which is also shown in figure 4, showing a scatter plot between fCover and the predicted iron content. The clear trend in figure 2 is removed and the number of pixels with large overestimation of the iron content decreased. Areas with fCover above 50% are not shown in figure 4.

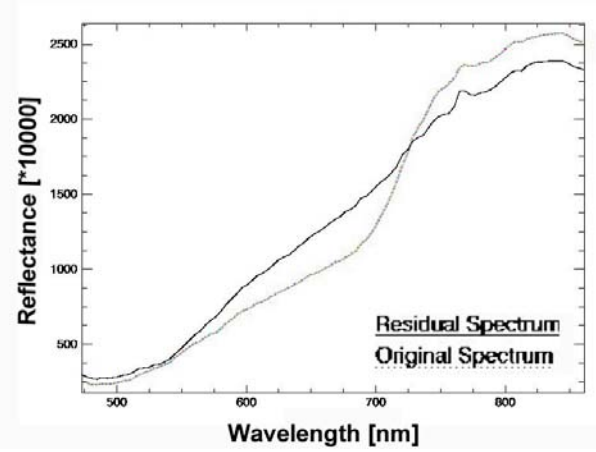

Figure 3: The spectral signature of a point with fCover = 0.32, before and after residual spectral unmixing. The influence of vegetation is removed.

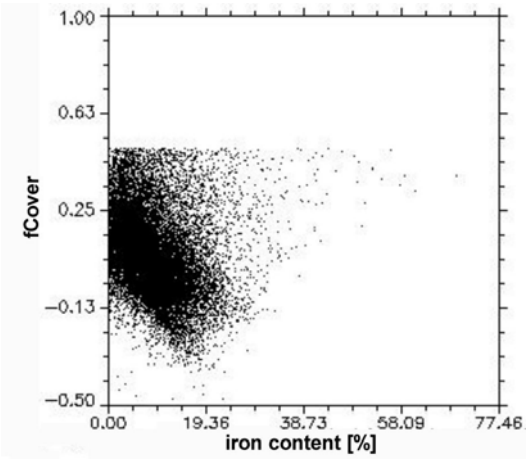

Figure 4: Scatterplot of the predicted total iron content and the fractional vegetation cover after residual spectral unmixing. The overestimation of the iron content at higher fCover is diminished.

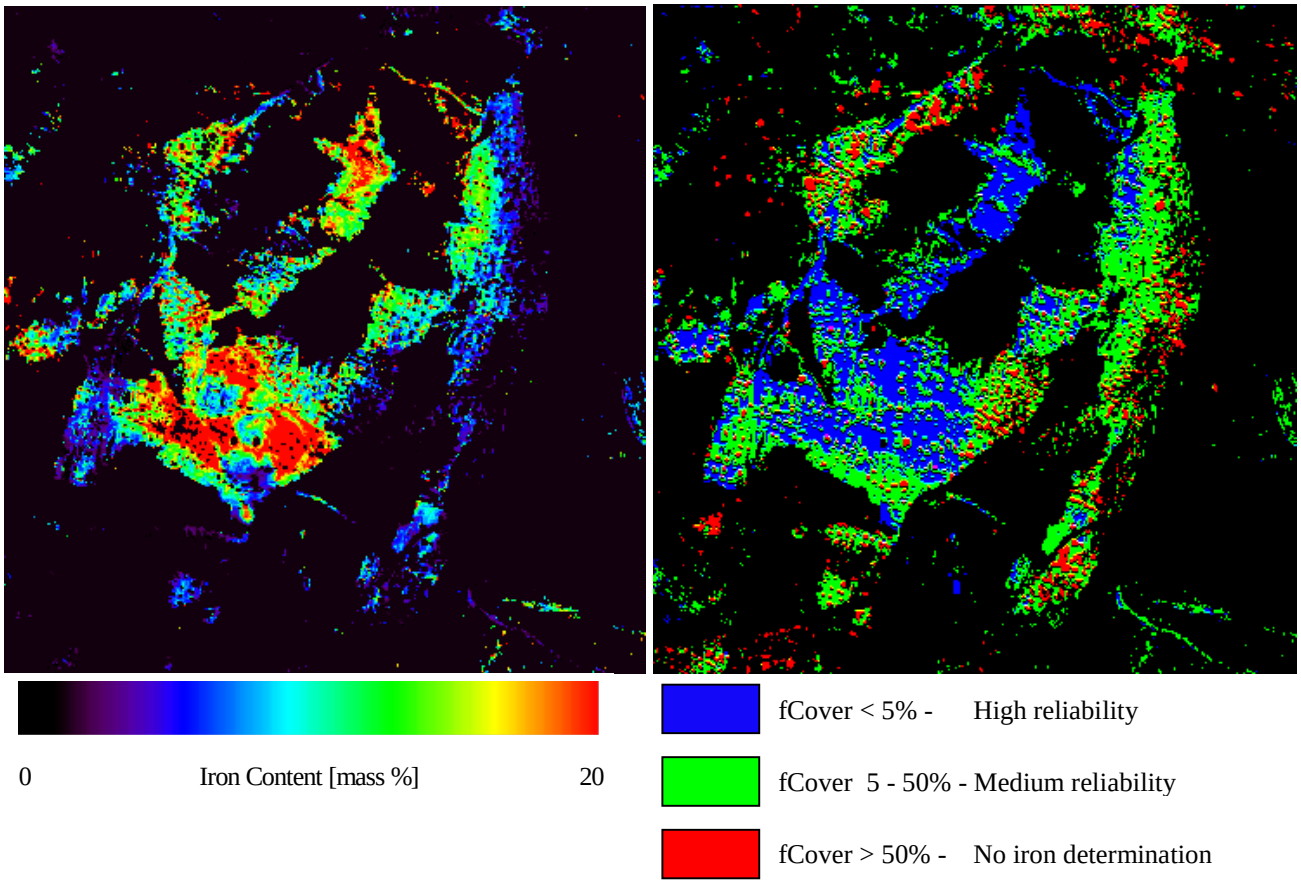

Figure 5: The left map shows the total iron content. It has to be used in combination with the accompanying map on the right, which shows the reliability of the data in the left map. Three reliability classes are defined, derived from the fractional olive cover.

Figure 5 shows the total iron content map of the study area, together with the map indicating the reliability of the used analysis.

The reliability of the iron quantification using the original smoothed spectra is considered the highest. The accuracy of the iron quantification on the residual unmixing procedure implies steps that are difficult to validate, making the final result in these areas less reliable. Especially the uncertain-

ties in the linear unmixing contribute to lower reliability of this procedure. The selection of representative end-members is a crucial step, but validation of unmixing results stays a complicated matter.

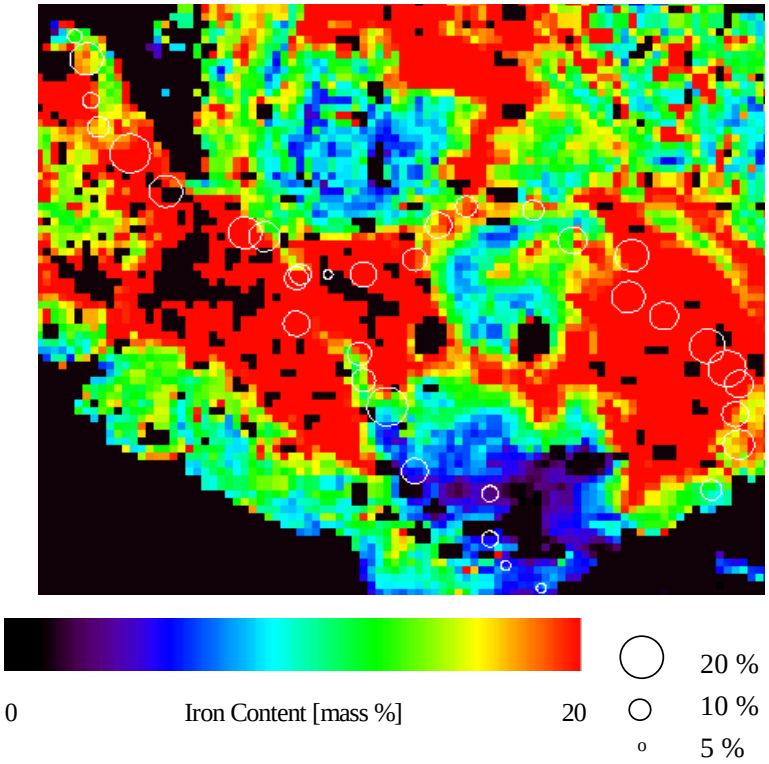

Figure 6: A subset of the study area where the total iron content is overlaid with the iron content derived from laboratory analysis.

Figure 6 shows the iron content of a subset of the study area, determined from the ROSIS images, combined with the iron concentration determined during the chemical analysis of the soil samples in the laboratory. It clearly shows the correspondence between the image derived iron content and the values determined with conventional techniques. When soil variables are mapped in the traditional way only a limited amount of points can be sampled, after which spatial interpolation techniques are used to get values for the entire study area. Imaging Spectroscopy delivers the tools for continuous mapping of soil variables, and with the described methodology the number of locations where interpolation techniques have to yield the required information can be diminished.

CONCLUSIONS

The influence of vegetation on the quantification of the soil total iron content is big. The magnitude of the error in iron prediction is dependent on the type of analysis that is used. In areas that are partially covered with olive, residual linear unmixing can remove the disrupting influence of the vegetation on the spectral response. However, due to complicated validation of the processing steps, this procedure is considered to yield less reliable results than the use of bare soil when predicting the total iron quantity. The residual unmixing procedure allows iron quantification in areas with partial vegetation cover.

ACKNOWLEDGEMENTS

We would like to thank Kees Klawer from Utrecht University to let us use the laboratory for the analysis of the soil samples. We acknowledge the use of the spectrometers from the Meetkundige Dienst Rijkswaterstaat and Radboud University Nijmegen.

References

1. Cohen, C.K., et al., The role of iron-deficiency stress responses in stimulating heavy-metal transport in plants. Plant Physiology, 1998. **116**: p. 1063-1072.
2. Liang, S., Quantitative Remote Sensing of Land Surfaces. 2004, New York: Wiley & Sons Inc. 481-482.
3. Escadafal, R., Remote Sensing of soil color: principles and applications. Remote Sensing Rev., 1993. **7**: p. 261-279.
4. Farrand, W.H., Identification and mapping of ferric oxide and oxyhydroxide minerals in imaging spectrometer data of Summitville, Colorado, USA and the surrounding San Juan Mountains. International Journal of Remote Sensing, 1997. **18**(7): p. 1543-1552.
5. Warell, J., Properties of the Hermean regolith: III. Disk-resolved vis-NIR reflectance spectra and implications for the abundance of iron. Icarus, 2003. **161**(2): p. 199-222.
6. Torrent, J., et al., Quantitative relationships between soil color and hematite content. Soil Science, 1983. **13**: p. 354-358.
7. Ben-Dor, E. and A. Banin, Diffuse reflectance spectra of smectite minerals in the near infrared and their relation to chemical composition. Sci. Geol. Bull, 1990. **43(2-4)**: p. 117-128.
8. Coyne, L.M., et al., Near-infrared correlation spectroscopy: quantifying iron and surface water in series of variably cation-exchanged montmorillonite clays. Spectroscopic characterization of Minerals and Their Surfaces, ACS Symp. Ser. 415, ed. L.M. Coyne, S.W.S. McKeever, and D.F. Blake. 1989, Washington. 407 - 429.
9. Bartholomeus, H.M., G.F. Epema, and M.E. Schaepman, Determining iron content in Mediterranean soils with fractional vegetation cover, using imaging spectroscopy. submitted, 2005: p. 15.
10. Buurman, P., Introduction to the interdisciplinary practical Sustainable Land Use in the Alora region, Spain. 1999, Landbouwuniversiteit Wageningen.
11. Raiswell, T., D.E. Canfield, and R.A. Berner, A comparison of iron extraction methods for the determination of degree of pyritisation and the recognition of iron limited pyrite formation. Chemical Geology, 1994. **11**: p. 101-110.
12. Berger, M., et al., The DAISEX campaigns in support of a future land-surface-processes mission. ESA Bull., 2001. **105**: p. 101-111.
13. Gege, P., et al. System analysis and performance of the new version of the imaging spectrometer ROSIS. In: 1st EARSeL Workshop on Imaging Spectroscopy. 1998. Remote Sensing Laboratories, University of Zurich, Switzerland.
14. Holzwarth, S., et al., HySens – DAIS 7915 / ROSIS Imaging Spectrometers at DLR. In: Proceedings of the3 rd Earsel Workshop on Imaging Spectroscopy, Herrsching, 13-16 may, 2003: p. 3-14.
15. Schläpfer, D. and R. Richter, Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: parametric orthorectification. International Journal of Remote Sensing, 2002. **23**(13): p. 2609-2630.
16. Green, A.A. and M.D. Craig, Analysis of aircraft spectrometer data with logarithmic residuals. In: Proceedings of the airborne imaging spectrometer data analysis workshop, Pasadena, USA., ed. G. Vane and A.F.H. Goetz. Vol. 85-41. 1985: NASA-JPL Publication. 111-119.
17. Smith, M.O., P.E. Johnstonn, and J.B. Adams, Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis. Journal of Geophysical research, 1985. **90**: p. 797-804.

18. Tompkins, S., et al., Optimization of endmembers for spectral mixture analysis. Remote Sensing of Environment, 1997. **59**(3): p. 472-489.
19. Green, A.A., et al., A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing, 1988. **26**(1): p. 65-74.
20. Van Der Meer, F. and S.M. De Jong, Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members. International Journal of Remote Sensing, 2000. **21**(15): p. 2781-2797.