# ENGEOMAP – A GEOLOGICAL MAPPING TOOL APPLIED TO THE ENMAP MISSION

*Christian Rogaß, Karl Segl, Christian Mielke, Yvonne Fuchs, and Hermann Kaufmann*

Helmholtz Centre Potsdam–GFZ German Research Centre for Geosciences, Section 1.4 Remote Sensing, Potsdam, Germany; E-mail: {christian.rogass / karl.segl / christian.mielke / yvonne.fuchs / hermann.kaufmann}(at)gfz-potsdam.de

## ABSTRACT

Hyperspectral imaging spectroscopy offers a broad range of spatial applications that are primarily based on the foregoing identification of surface cover materials. In this context, the future hyperspectral sensor EnMAP will provide a new standard of highly qualitative imaging spectroscopy data from space that enables spatiotemporal monitoring of surface materials. The high *SNR* of EnMAP offers the possibility to differentiate and to identify minerals that are showing characteristic absorption features as a 30 m × 30 m spatial mixture in the visible, the near infrared and the short wave infrared range (0.4 - 2.5 µm). For this purpose, spectral mixture analysis (SMA) approaches are traditionally used. However, these approaches lack in transferability, repeatability and inclusion of sensor characteristics. Additionally, they rely on image-based and randomly detected endmembers as well as on *in situ* or laboratory spectra that are not spatially stable in case of an image-based extraction and are assumed to be spectrally pure. In this work, a new framework is proposed that addresses these limitations considering the EnMAP sensor characteristics. It is named EnMAP Geological Mapper - EnGeoMAP. It consists of several new and adapted approaches to identify spectrally homogeneous regions. In parallel, minerals are identified and semi-quantified by a sensor-related and knowledge-based fitting approach. Supplementary outputs are abundance, classification, homogeneity and uncertainty maps. First results show that the proposed approach offers 100% repeatability and gains an identification error for minerals of about 2% on average for different studies. In this work, an approach is proposed that aims on spectroscopic mineral modelling by image synthesis that might be applied for geological mapping.

## INTRODUCTION

Remote sensing of soils and geology often relies on approaches that directly identify minerals within hyperspectral images. To achieve this, unknown image spectra are statistically compared with known library or *in situ* spectra. Field samples are often additionally analysed by X-Ray diffractometry (XRD) and by fluorescence spectroscopy (XRF) for identification and quantification. Based on geochemical and spectroscopic analyses the absorptions of different minerals are identified or modelled and defined as diagnostic spectral features. These features are to some extent unique for each mineral. Additionally, analysed minerals are assumed to be pure or spectral impacts of insignificant fractions of elements on mineral compounds are neglected. However, minerals often form partial solid solutions, e.g. pyroxenes. Rocks that are built by rock forming minerals might be unique in texture and spatial distribution according to their geological and petrologic evolution. Hence, *in situ*, airborne and spaceborne acquired spectra show rather unique region-related mineral mixtures than pure minerals. This makes it more difficult to identify observed minerals and their fractions within one pixel. Since mineral identifications are frequently conducted in mountainous regions shadows aggravate any kind of identification due to the decrease of reflected incident radiation. Contemporary hyperspectral sensors may considerably differ in their sensing principle, spectral and spatial resolution. This leads to sensor-specific scaling phenomena between the sensors and *in situ* or laboratory spectroscopy.

Within the framework of the EnMAP project (1) an approach was developed – named EnMAP Geological Mapper (EnGeoMAP) – that aims on the reduction of the previously described signal impacts. The core algorithm of the EnGeoMAP is similar to the broadly accepted Tetracorder (2)

that represents a broadly accepted, knowledge-based expert system for the spectroscopic differentiation of minerals. However, it dynamically and iteratively utilises properties of the inspected acquisition and its sensor. Characteristics such as sensing geometry, spatial and spectral resolution and terrain are considered. It consists of a multistep algorithm and enables soil and geological applications such as mineral mapping, alteration zone detection, mine waste characterisation and many more. Additionally, quality flags for each inspected pixel are given. These can be incorporated in further analyses, e.g. classifications, or in succeeding iterations. The approach was tested in the Makhtesh Ramon of the Negev Desert in Israel. Here, geology has been studied for decades. One artificial EnMAP scene was synthesised on the basis of one real hyperspectral airborne scene to objectively evaluate the EnGeoMAP identification results of a spaceborne hyperspectral acquisition.

## MATERIALS AND METHODS

The Makhtesh Ramon of the Negev desert in Israel (Fig. 1) was selected as case study region. The Makhtesh consists of different mineral compounds that are known and useful for testing geologically related remote sensing algorithms (3). In preparation for the EnMAP mission an atmospherically corrected hyperspectral AISA DUAL (4) scene was used as a basis to simulate an EnMAP scene with the EnMAP-End-to-End-Simulator – EETES (5). This scene was acquired on the 15.03.2004 at 30.4°N / 34.5°E incorporating a sun azimuth of 210° and a sun elevation of 54°.

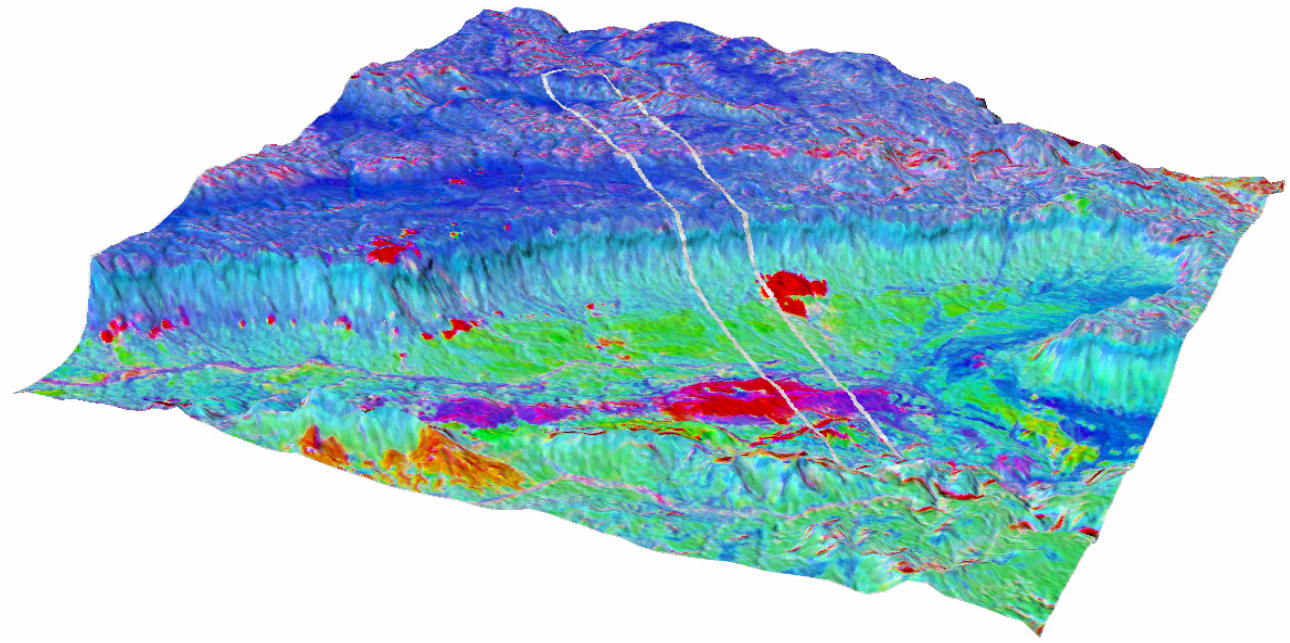

*Figure 1: Case study region Makhtesh Ramon here figured as hill shaded 3D false colour composite that is grey overlayed by the extent of the hyperspectral AISA DUAL acquisition (RGB = Landsat TM (GSD 30 m) mineral ratios 5/7 (clays), 5/4 (ferrous iron) and 3/1 (iron oxides); 3D = ASTER DEM (is a product of METI and NASA, GSD 15 m)).*

During the EETES simulation many sensor parameters from the manufacturer were considered, such as the orbit parameters, Point Spread Function (*PSF*) for each detector, spectral and radiometric responses. In consequence, the simulated EnMAP scene consisted of 244 bands ranging from 400 to 2500 nm and incorporating a ground sampling distance (*GSD*) of 30 m. In this work, we focused on the spectral range from 2 to 2.5 µm of the Short Wave Infrared (SWIR) that includes most significant diagnostic spectral features of minerals.

In addition, missing illumination caused by shadow casting objects such as mountains can be nearly linearly continuum normalised (compare Figure 2 that shows an average deviation from linearity of about 0.8%) only in the SWIR range.

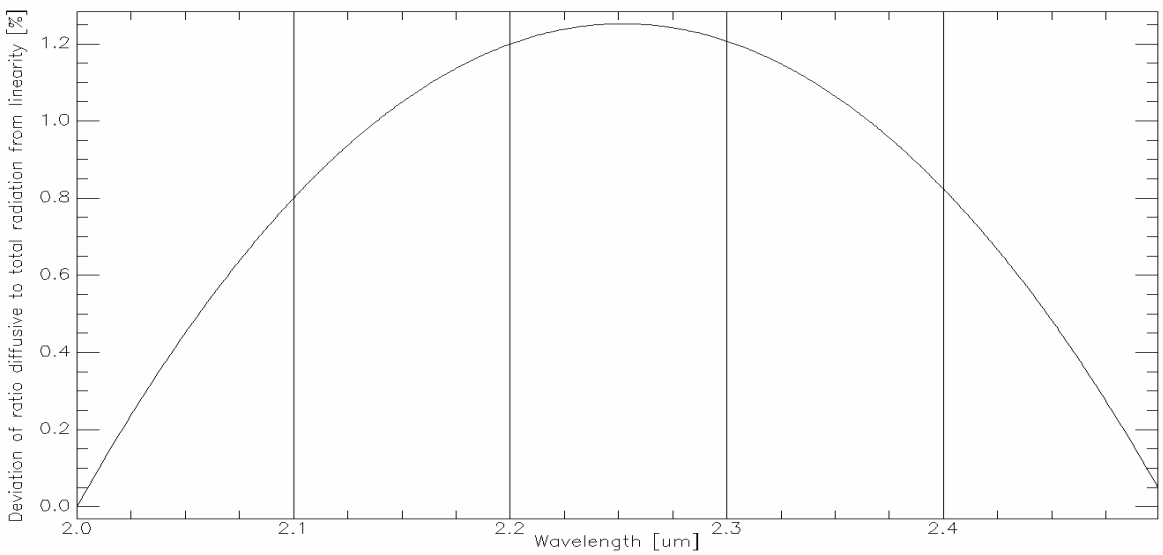


*Figure 2: Deviation of the ratio between diffusive and total radiation from linearity.*

Due to inadequate research on the impact of shadow and its removal approaches on mineral identification techniques, the scene was not corrected for shadows. However, the analysis of shadowed regions implies a reduction of the identification accuracy, since the Signal-to-Noise-Ratio (*SNR*) is significantly lower than in directly illuminated regions. Although this effect is broadly accepted, most geological mappers (6) do not fully consider the relationship between *SNR* and identification accuracy. Most of them directly compare known library spectra with unknown spectra directly assessed from the image as endmembers (7). In this work, we rely on the USGS spectral library (8) and the feature descriptions of the Tetracorder (2). However, this algorithm can be considered as a knowledge-based expert system to directly identify spectra of hyperspectral acquisitions. The Tetracorder has proven its applicability in the past, but provides only limited capabilities of analysing mineral compounds by modelling abundances of exclusive features (2). However, this algorithm and other widely used alogorithms (6) only consider the spectral characteristics of the sensor but not the spatial impact of the GSD and the PSF on the distribution and the abundances of spectra.

In this work, we propose a sensor-related approach that fully incorporates sensor characteristics such as the Spectral Response Function (*SRF*) and the *PSF*. With increasing sensor *GSD* the likelihood of having spectrally pure material in one pixel decreases. To avoid confusion, the material compound of each pixel is here considered to be a mixture and spectrally homogeneous regions are considered to be flat fields composed of basic mixtures. This definition is independent of the sensor and the acquisition geometry. Furthermore, it was assumed that in spectrally homogeneous regions (flat fields) pixels are linear mixtures also incorporating adjacent pixels within the effective range of the *PSF* for this detector and wavelength. The relevant range used for this work was 99 % of the *PSF*'s volume. Additionally, it was assumed that the *PSF* of different detectors are similar shaped and per scene constant. In this case, each pixel of a flat field is an isotropic mixture of itself and its neighbourhood. Nonlinear effects in flat fields only exist, if *BRDF* effects occur and, hence, the spatial extent of inspected neighbourhood should be rather narrow.

The EnGeoMAP consists of three modules – the FeatureLUT, the Basic Mixture identification and the Mixture analysis that are sketched in Fig. 3 and described in more detail in the following.

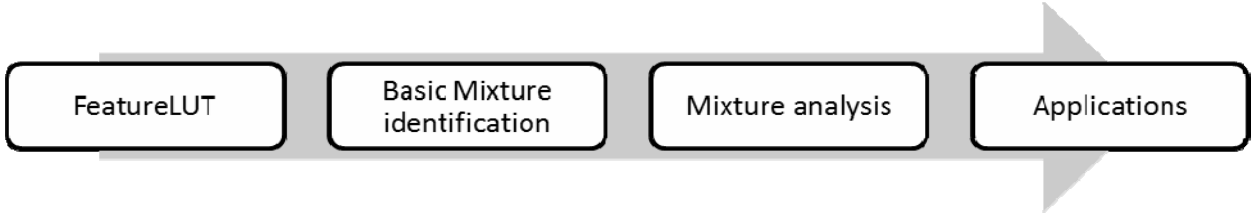


*Figure 3: Workflow and relation of the modules of EnGeoMAP.*

**Module 1 – FeatureLUT**

The Look-Up-Table (LUT) of the EnGeoMAP named FeatureLUT consists of more than 100 mineral spectra and their feature and fitting descriptions from USGS. It is similar to the Tetracorder (2) but extended with additional entries for chemical formulas, mineralisation type, alteration type etc. To use the library and the feature descriptions with different sensors, all criteria and spectra were resampled to 1 nm resolution. In case spectra and criteria had a lower spectral resolution than 1 nm, Hermite Splines were used for interpolation. After resampling to 1 nm, a re-usable FeatureLUT was created that is still sensor-independent. This sensor-independent FeatureLUT is then resampled to the sensor. The spectral resampling is performed by spectral deconvolution (9). Knowledge-based fitting thresholds are also adapted in the process of resampling. This is necessary because predefined thresholds (as in Tetracorder) depend on incorporated sensors leading to misidentifications of spectra acquired by sensors with a higher spectral and spatial resolution and a better *SNR*. After resampling, the FeatureLUT is sensor-dependent, re-usable for this sensor unless its characteristics have changed, and it serves as a basis for next processing steps of the EnGeoMAP.

**Module 2 – Basic Mixture identification**

This module consists of three steps – the flat field detection, the mineral identification (core of the EnGeoMAP) and a Bounded Value Least Squares (*BVLS*) unmixing.

The flatfield detection is based on the assumption that in spectrally homogeneous regions mixtures are related to the PSF. In this process a moving window of an adapted size that relates to the 99 % volume threshold of the mean sensor PSF is used to locally compute the uncentred Pearson correlation coefficient between the spectrum of the centre pixel of the window and the spectra of all the neighbours within this window. To suppress albedo effects, the continuum of each spectrum is removed by normalising with its Delaunay approximated convex hull. In the process of continuum removal, concave curve shapes are preserved that correspond to absorptions. If all fits pass a predefined fitting threshold (by default 0.99), the pixel is binary marked as flat field pixel. This is performed for all pixels in the scene and results are stored in a binary map where all flat field pixels are marked.

After detecting flat field locations the mineral identification is carried out. For this, each spectrum of the flat field is fitted towards all library spectra of the FeatureLUT within their specific diagnostic features described by Tetracorder. All fits that pass the sensor-adapted thresholds, similar to Tetracorder, are stored in a local pixel-related list. After this, a *BVLS* unmixing is performed for this pixel that excludes all identified spectra that do not pass an unmixing threshold (by default 5 %, but *SNR* dependent) to remove outliers that are too noisy and not significantly abundant. All remaining, identified spectra are stored in a global list and all identification results for this pixel are rejected. Then, the next pixel of the flat field is considered and the global list is updated. This is performed until all pixels of the flat field have been inspected. Consequently, a global list of matching spectra is created that is directly used in the next step.

**Module 3 – Mixture analysis**

In this step, all pixels are linearly unmixed on the basis of the global list by the *BVLS*. Again, outliers are removed by applying an abundance-related threshold (also 5 % minimum abundance by default, but *SNR* dependent). After this, the image is synthesised by using previously estimated abundances of identified mineral spectra.

This enables a model image to be generated that is directly comparable with the continuum removed real image. Consequently, each pixel gets an individual model error that helps to distinguish regions where spectra were accurately identified from problematic regions such as shadow regions. Additionally, a next iteration can be applied starting with module 2 to exclude these areas in advance.

As a result, each individual pixel provides wavelength-dependent information of mineral abundances, error budget and flat field potential. This is then directly applicable to hydrothermal alteration mapping in a next step such as spatial pattern analyses.

## RESULTS

The potential of the EnGeoMAP is here exemplarily demonstrated for the analysis of one hyperspectral, synthesised EnMAP scene. The evaluation of the results is based on the assumption that only a correct identification of diagnostic features and a correct estimation of abundances of minerals provide low deviations between modelled and real image spectra. The bands that encompass dominating mineral absorption features should be spatially and spectrally equivalent. This condition was mostly fulfilled for given examples (Figure 3).

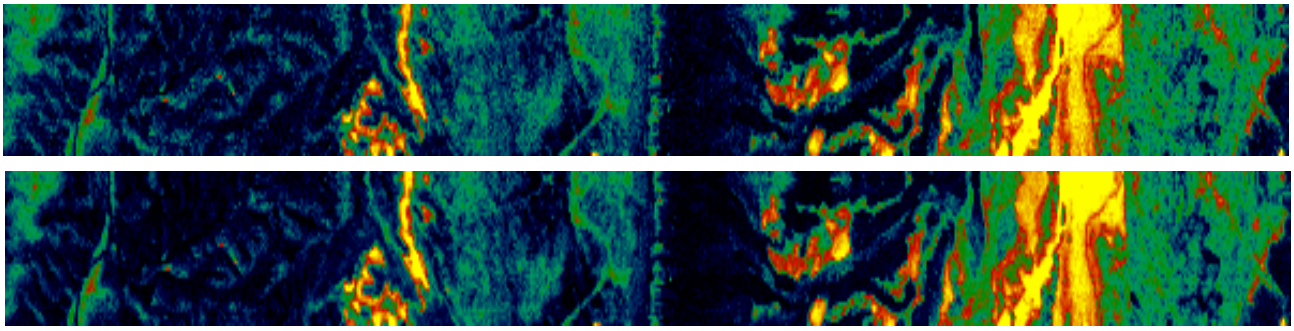

*Figure 4: Inversed colour composite for the clay absorption at 2.2 µm of the real image (top) and the model image (bottom).*

For the 2.2 µm band shown in Fig. 4 a deviation between the model and the real image of about 0.5% was achieved that is close to the overall accuracy of 0.8% on average for all bands. However, the accuracy of EnGeoMAP is decreased up to 20 times for the whole spectral range in low *SNR* regions. Comparing the mean ratio of the diffusive to total radiation within the spectral range between 2 and 2.5 µm (about 10%) with the accuracy decreasing rate (about 20 times) for low *SNR* regions as in shadows clearly shows a strong relationship between the accuracy and the *SNR*. This is also shown in Fig. 5 for three plots representing three different extreme *SNR* scenarios (Plot 1 – average SNR, Plot 2 – low SNR, Plot 3 – high SNR).

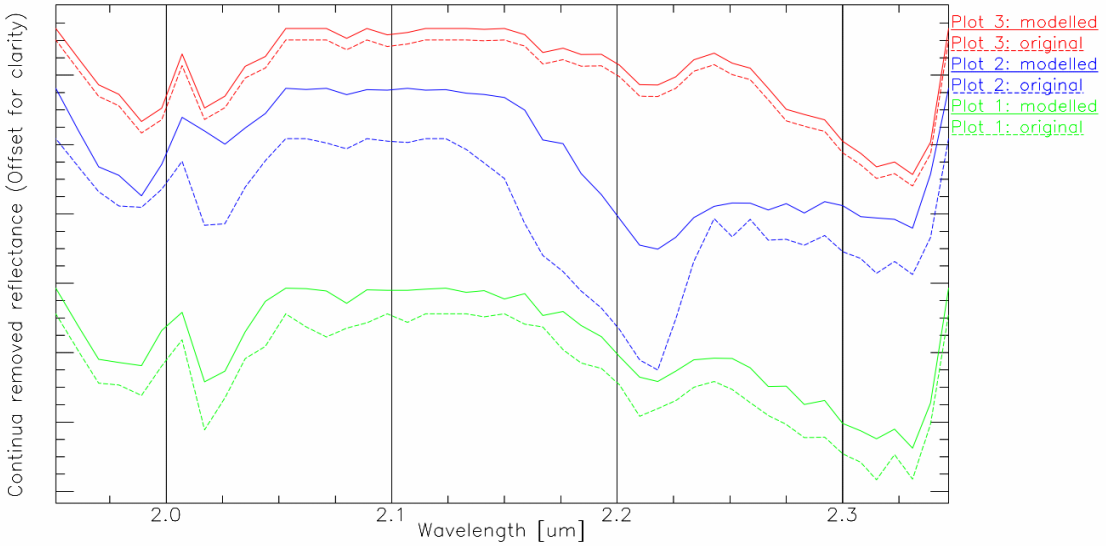


*Figure 5: Sample plots of continua removed modelled (fully analysed) spectra vs. real spectra – Plot 1: 60% Carbonate, 10% Epidote, 30% Clay and a model error of 4% – Plot 2: 30% Carbonate, 70% Clay and a model error of 15% - Plot 3: 75% Carbonate, 25% Clay and a model error of 0%.*

Considering only the results for these extreme regions reveals the range of potential uncertainties in assessing mineral contents, although the assessment is superimposed by the evaluation of non-relevant spectral regions. This is exemplarily shown in Fig. 6 depicting the maximum error along the spectral dimension for each analysed pixel and its likelihood for a low albedo.

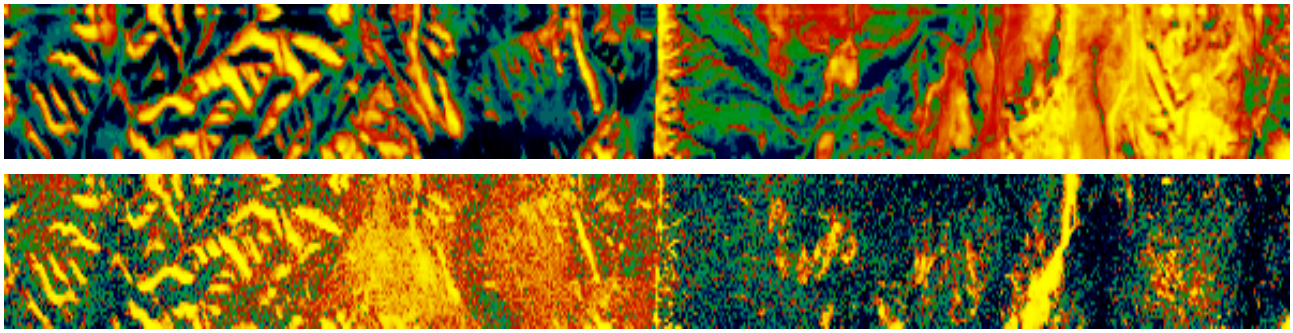

*Figure 6: Upper two quartiles of shadow and dark material abundances (top) and error budget for this scene (bottom, maximum 16% and minimum 0%) having a common spatial correlation of 88% on average.*

In any case, analyses in low *SNR* regions as in shadows should be marked spatially and error budget-related to avoid relying on average error budgets for the whole scene. A positive side effect is the potential of having weights for succeeding analyses such as classifications as given in Fig. 7 that illustrates abundance dominating minerals.

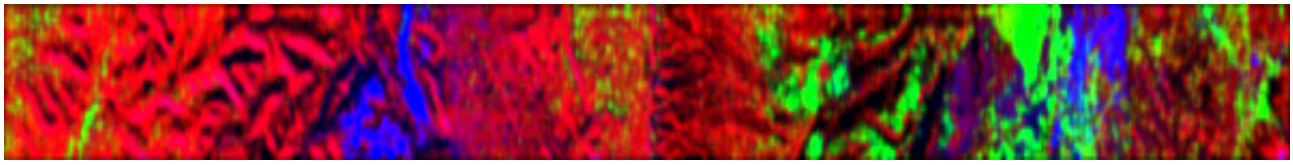

*Figure 7: False coloured RGB abundance composite of dominating minerals (Red – Carbonates, Blue – Epidote, Green – Clay minerals).*

In all, EnGeoMAP achieves an identification accuracy of about 99% on average and of about 98% on average for the maximum error budgets along the spectral dimension.

Uncertainties in the BVLS unmixing and in the identification of minerals in low SNR regions have remained so that pixel-based error budgets can be used to exclude erroneous analyses from further processing such as classification.

## CONCLUSIONS

EnGeoMAP achieved high identification accuracy for this case study region. It is completely unsupervised, 100% repeatable since no random statistics are used, platform-independent and will be freely available as soon as it is implemented in the free EnMAP software named EnMAP box. Additionally, spectral and spatially homogeneity maps are provided that might be useful for other processing such as segmentation. Currently, more hyperspectral scenes are acquired in Southern Africa, Mongolia and Spain that will be used to verify and further improve the proposed EnGeoMAP approach. Additionally, XRD and XRF data will be analysed to evaluate both the reliability of the reference spectral library and the mapping results.

## REFERENCES

1 Kaufmann H, K Segl, L Guanter, S Hofer, K P Foerster, T Stuffler, A Mueller, R Richter, H Bach, P Hostert & C Chlebek, 2008. Environmental Mapping and Analysis Program (EnMAP) - Recent advances and status. In: Proceedings International Geoscience and Remote Sensing Symposium (IGARSS), Boston, MA, 7-11 July, IV-109-IV-112

2 Clark R N, G A Swayze, K E Livo, R F Kokaly, S J Sutley, J B Dalton, R R McDougal & C A Gent, 2003. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. Journal of Geophysical Research, 108(5131): 1-44

3 Anker Y, E Ben Dor, E Zelikman, A Karnieli & E Mazor, 2009. Makhtesh Ramon, a super site for calibration and validation of is sensors. In: Proc. 6th EARSeL SIG Workshop on Imaging Spectroscopy (Ramat Aviv, Tel Aviv, Israel, 16-19 March)

4 Spectral Imaging Ltd. Aisa Dual, 2nd Version. http://www.specim.fi/media/aisa-datasheets/dual_datasheet_ver1-2012.pdf (last date accessed: 5 June 2012).

5 Segl K, L Guanter, C Rogass, T Kuester, S Roessner, H Kaufmann, B Sang, V Mogulsky & S Hofer, 2012. EeteS - The EnMAP End-to-End Simulation Tool. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 5(2): 522-530

6 van der Meer F, 2004. Analysis of spectral absorption features in hyperspectral imagery. International Journal of Applied Earth Observation and Geoinformation, 5(1): 55-68

7 Boardman J W (1994). Geometric mixture analysis of imaging spectrometery data. In: Proc. International Geoscience and Remote Sensing Symposium (IGARSS), Pasadena, CA, 8-12 August, IV-2369-IV-2371

8 Clark R N, G A Swayze, R Wise, E Livo, T Hoefen, R Kokaly & S J Sutley, 2007. USGS digital spectral library splib06a. U.S. Geological Survey, Digital Data Series 231 (last date accessed: 23 August 2012)

9 Guanter L, K Segl, B Sang, L Alonso, H Kaufmann & J Moreno, 2009. Scene-based spectral calibration assessment of high spectral resolution imaging spectrometers. Optics Express, 17(14): 11594-11606