DERIVATION OF SYNTHETIC ENDMEMBERS FOR LINEAR UNMIXING TO IMPROVE PARAMETER ESTIMATION FOR SOIL EROSION MODELLING IN AGRICULTURAL ECOSYSTEMS

Heike Gerighausen, Erik Borg

German Aerospace Centre, German Remote Sensing Data Centre, Neustrelitz, Germany; heike.gerighausen@dlr.de, erik.borg@dlr.de

ABSTRACT

Physically based erosion models have been widely accepted in soil erosion risk assessment in the past years but their application is still restricted due to a large disparity between required model input parameters and data availability. Hyperspectral imaging may offer an alternative to solve that problem providing high spatial and spectral information about earth surface materials. A major problem in image interpretation is the heterogeneous composition of each individual pixel, in the case of agricultural ecosystems a mixture of soil and vegetation signal, but also moisture conditions and illumination have a strong influence. This issue has been addressed with unmixing techniques (vii) which take individual pixel components, endmembers, into account.

In this paper, we present a new algorithm which allows the derivation of synthetic spectral endmembers (SY-SEM) of soil. A systematic analysis of lab spectra from more than 100 soil samples measured at different soil moisture and illumination angles revealed stable features independent of soil moisture and illumination in SWIR I and II. Results were confirmed by an analysis of corresponding field spectra. Starting from these stable wavelength regions absolute reflectance values of all spectra were analysed adopting a linear regression analysis. Results showed coefficients of determination (R^2) generally above 0.8, in the range of 1000 to 2500nm above 0.95. Based on these results, SY-SEM are derived from self-generated look-up tables containing regression parameters, and atmospherically corrected image data. In contrast to existing approaches, SY-SEM procedure is computed pixelwise over a pre-classified image data set.

INTRODUCTION

Soil erosion has remained a serious issue in agricultural ecosystems causing severe ecological and economical consequences. To approach this problem, a quantitative assessment of erosion risk and soil loss is essential. Physically based erosion models such as EUROSEM (x, xi) or EROSION-3D (xiv, xv) became widely accepted in mesoscale to local-scale modelling. However, they require numerous, high-quality input parameters in terms of spatial and temporal precision. In-situ point measurements and existing data catalogues cannot meet these requirements. In contrast, hyperspectral data offers a high density of spatial and spectral information which is an alternative as model input data source, as has been demonstrated by different authors (e.g. iv, v, ix).

Despite the comparatively high spatial resolution of hyperspectral airborne data (4.5m at approx. 1.5km flight altitude for HyMap) spectral signals rarely represent pure objects but are a mixture of different surface materials. Because of the complexity of the physical relationship between spectral reflectance values and material properties, it is inevitable to extract the signal of the individual pixel components rather than interpreting the mixed signal in order to derive relevant parameters in erosion modelling. Spectral mixture analysis (SMA) (i, ii) has become an outstanding method which addresses the problem of mixed pixels. It assumes that the spectral signal is an assemblage of elementary constituent spectra, called endmembers. In this respect, the pixel to pixel variability in a scene can be explained with varying proportions of spectral endmembers.

Endmember collection and calibration is an important item in the application of SMA. Their quality significantly controls the quality of the output result. The extraction of endmembers from lab and

field spectra has drawbacks since they are rarely acquired under the same conditions as airborne or satellite data. The increasing availability and accessibility of hyperspectral data in recent years have been accompanied by a boost in the number of different automated techniques for endmember extraction from image data. Such techniques are Pixel Purity Index (PPI) (iii), fast autonomous spectral endmember determination in hyperspectral data (N-FINDR) (xi), Iterative Error Analysis (IEA) (xii), Sequential Maximum Angle Convex Cone (SMACC) (viii), and Automated Morphological Endmember Extraction (AMEE) (xiii). These procedures analyse image data either without incorporating information on spatially adjacent data or by integrating spatial and spectral information to extract a predetermined number of endmembers. In both cases, they do not take effects of moisture or illumination on individual endmember characteristics into consideration. Furthermore, a very high computational complexity is inherent in these extraction methods. Therefore, a fast assessment of endmember spectra on the basis of special spectral features with regard to individual endmember constitution could support the elaborate extraction procedure, in particular, of very high dimensional datasets.

The objective of this paper is to discuss a method for deriving synthetic spectral endmembers (SYSEM) from hyperspectral images on the basis of self-generated look-up tables. Endmembers are derived pixel per pixel based on a preclassified image instead of a limited number of endmembers directly retrieved from image pixel spectra which are presumed to represent the entire endmember range.

In the first section of the paper, the study area, soil reference data base, and reflectance measurements are introduced. This is followed by a presentation of the methodical concept and a discussion of results from application to the presented data base. The paper closes with a summary and conclusions.

METHODS

Study Area

The study area is located in the northeast of Germany, in the centre of Mecklenburg-Western Pomerania, and is part of DEMMIN (Durable Environmental Multidisciplinary Monitoring Information Network), a test site for calibration and validation of value added products from remote sensing. The site is an intensively used agricultural ecosystem comprising five farms which are combined within a farming association with a total area of about 25.000ha (Figure 1). Single fields are very large in size averaging about 80ha. The main crops grown in this area are winter wheat, winter barley, winter rape, maize and sugar beet.

The area of DEMMIN is a typical Pleistocene landscape with hilly, loamy to clayey moraines, sandy plains and peaty floodplains (vi). The altitudinal range is about 80m. In general, precipitation per year is moderate at 500 to 600mm. But torrential rainfall triggers large-scale soil erosion events and is promoted by unfavourable cultivation.

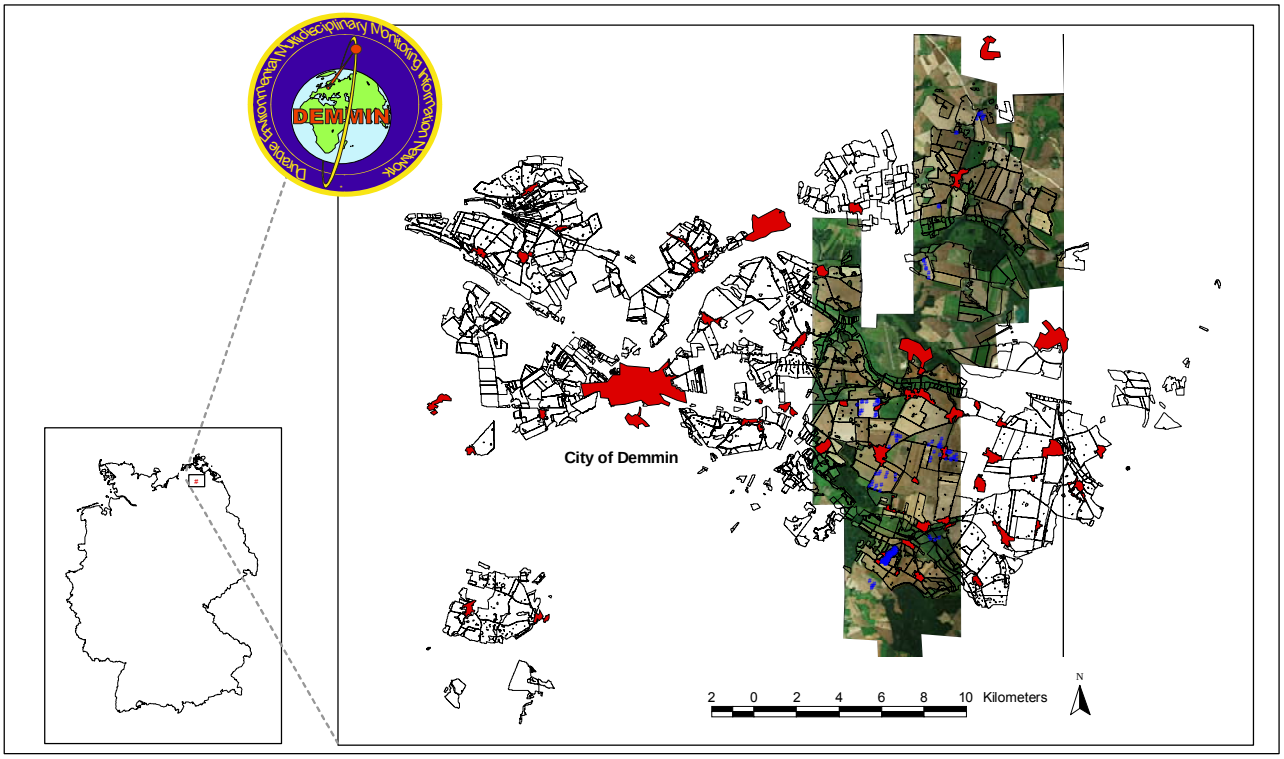

Figure 1: Location of the DEMMIN test site in the northeast of Germany comprising 25.000 ha of agricultural fields. Focus of the study are the HyMap flight tracks (Acquisition date: 09.08.2004, true colour composite). Blue signs mark sampling points.

Soil reference data base

One hundred-thirty-five soil samples of approximately 1 litre were collected from the top 5cm of soil from the agricultural fields in DEMMIN. Samples sites were selected using existing soil maps and a digital terrain model to represent the scope of soil types as complete as possible.

The soil samples were analysed by ZALF (Centre for Agricultural Landscape Research, Central Laboratory) in Müncheberg to determine the particle size distribution, organic matter content, total iron, carbon content and pH-value. Table 1 shows the statistics for all soil samples. Characteristic soil texture classes are sand, loamy sand and loamy to silty clay. 13% of the samples have clay contents higher than 20% with a maximum clay content of 33%. Organic matter (OM) contents range from 0.5 to 1% for 50% of the samples. 30% have OM contents of 1 to 1.5% reaching a maximum of 2.48%. Iron contents vary from 0.46 to 2.08%. The mean pH-value as indicator of soil acidity is 6.92. 60% of the samples show pH-values between 7 to 8, stating light alkaline conditions.

Table 1: Soil physicochemical characteristics from DEMMIN test site.
Particle size fractions: clay 0-2µm, silt 2-63µm, sand 63-2000µm.

	Sand (%)	Silt (%)	Clay (%)	OM (%)	Iron (%)	pH	CO3-C (%)
Mean	61.59	25.72	12.69	1.03	1.04	6.92	0.13
Min	38.02	6.19	3.65	0.43	0.46	4.43	0.00
Max	89.96	39.57	32.95	2.48	2.08	7.67	1.08
Std.	10.31	5.89	6.53	0.39	0.38	0.61	0.21

Reflectance measurements

Spectral measurements were acquired using an ASD spectroradiometer FieldSpec Pro covering the full spectral range from 350 to 2500nm. All measurements were taken with bare fibre and a 25° FOV whereas a measurement at each sampling point or soil probe was averaged out of ten measurements with a system integration time of 50 per single collected spectrum. A standardised spectralon panel was used as reflectance reference.

For lab measurements soil samples were grounded with a mortar and passed through a 2mm sieve to reduce anisotropic light scattering. To standardize the moisture level they were oven-dried for 48h at 105°C. For each soil sample, a series of four different moisture contents was prepared that is oven-dry, 7.2M.-%, 10M.-% and 13M.-%. Soil wetting was accomplished using a spray bottle, adding the according mass fraction of water. A balance of 10mg accuracy was used for weighing. To guarantee homogenous water dispersion of the samples, soil was mixed with a stirring staff immediately after wetting. One soil sample series was prepared at a time for measuring to minimise the error due to surface drying-up. A 1000 W tungsten lamp was used as light source to illuminate the target at an incident angle of 30°. Oven-dried soil samples were additionally measured at incident angles of 45° and 60°. An illustration of the experimental set up in the laboratory as well as sample conditions in the lab and field is shown in Figure 2.

Field reflectance measurements were carried out in May and September 2006, shortly after sowings to benefit from that short time period of bare soil on agricultural field parcels. Nevertheless, in some cases, seedlings had to be removed.

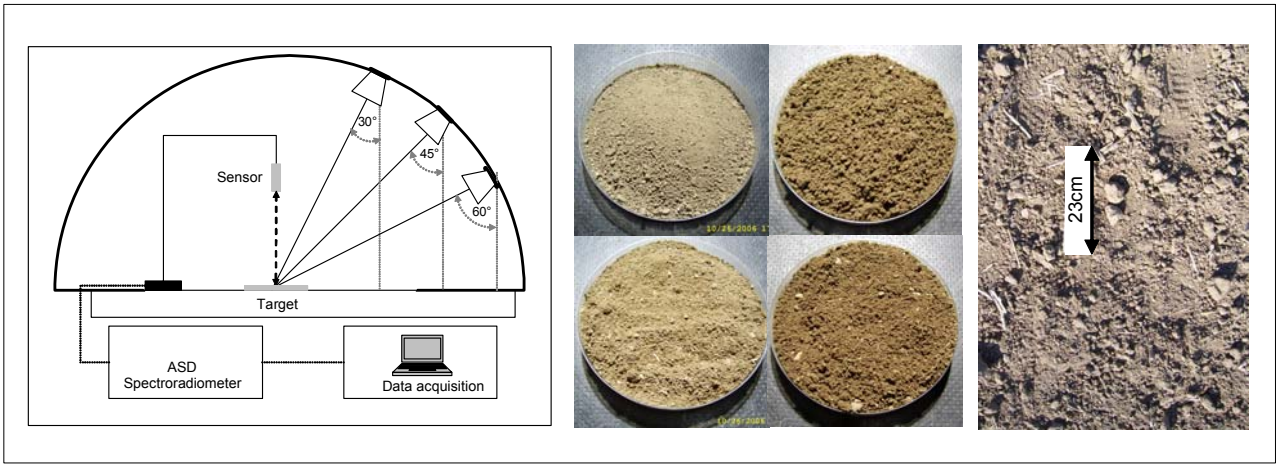

Figure 2: Laboratory configuration for spectral measurements and sample conditions during measurement in the lab (middle) and in the field (right).

RESULTS

Methodical Concept

The presented approach for synthetic endmember derivation is divided into two parts. The workflow is shown in Figure 3. The first part consists of a systematic pre-analysis of lab and field spectra to identify stable spectral features which are independent of soil physicochemical characteristics, soil moisture and sun incident angle. The second part concentrates on the derivation of synthetic endmembers from realistic mixed pixel image spectra. It is established on the results of the preliminary study and remote sensing image data.

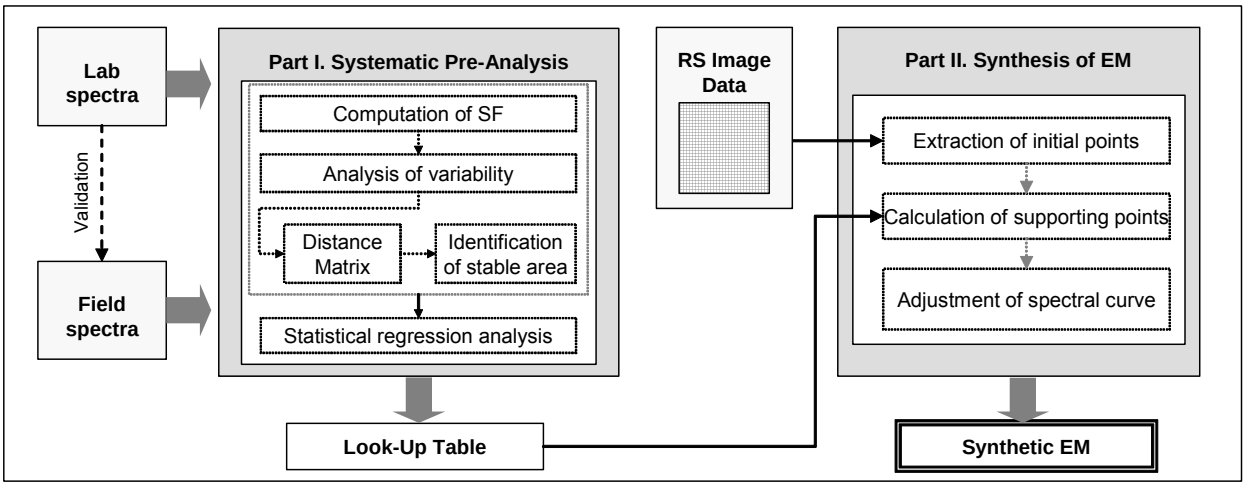

Figure 3: General workflow for synthetic endmember derivation.

The systematic analysis of lab and field spectra is carried out with the help of two-dimensional matrices. Features are computed stepwise for all possible wavelength combinations of a spectrum and subsequently stored in a two-dimensional matrix where individual pixel coordinates correspond with the according input wavelengths (Figure 4). The features selected for calculation are, for example, ratio, absolute or normalised difference. The resulting matrices are stacked and the range is computed pixelwise. Simultaneously, a wavelength distance vector matrix is established to indicate the total offset between a pair of wavelengths in the matrix. This way a reasonable offset for stable feature extraction can be taken into account. For the study, it was set to a minimum of 50nm. The identification of stable wavelength features is followed by a statistical analysis between reflectance values of potential stable wavelengths features and all other reflectance values to predict additional supporting points for SY-SEM derivation. The results of the analysis, that are the coefficients of determination (square of PEARSON correlation coefficient, R^2), slope and bias, are finally stored in look-up tables.

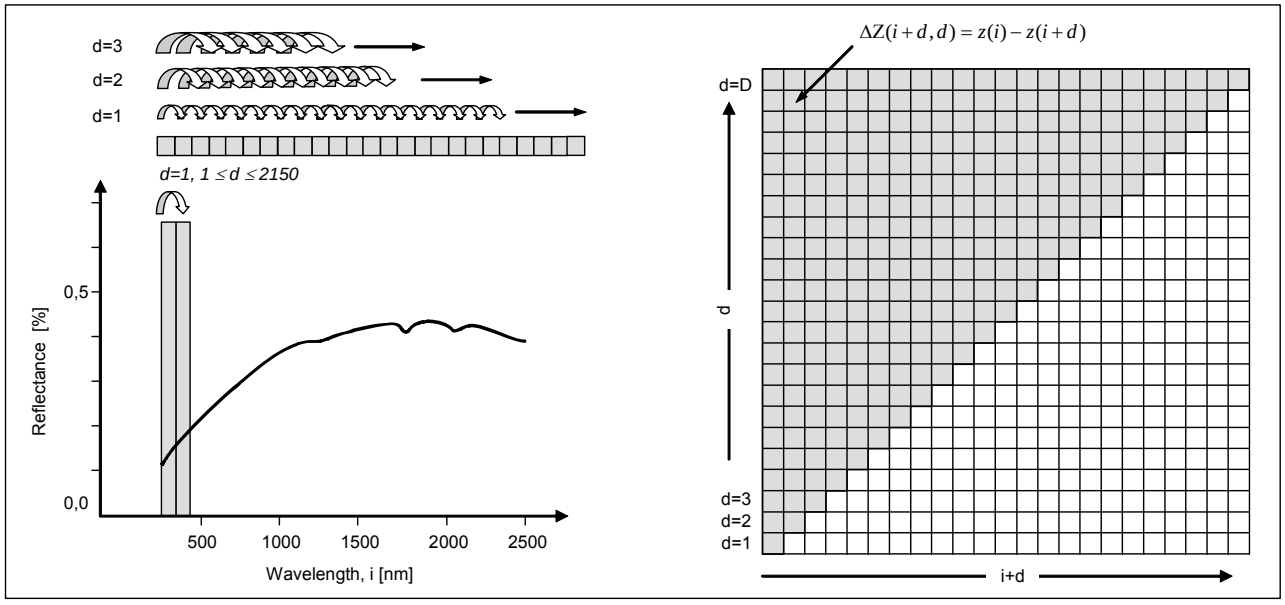

Figure 4: Generation of a triangular matrix from a spectral curve.

This first part of the procedure has to be done only once per endmember within a regional approach. Therefore, if study environments do not change significantly, look-up tables can repeatedly be applied.

These look-up tables together with the remote sensing image data build up the knowledge base of the second part, which is the derivation of SY-SEM itself. According to the results of part one, reflectance values of initial points for endmember synthesis are extracted from remote sensing image data. Starting from these points, additional supporting points are computed deploying look-up tables. The last step is a curve adjustment at wavelengths of low correlation.

Analysis results

A main objective of the systematic analysis of lab and field spectra in the preparation of the study was the identification of wavelength combinations independent of moisture and illumination as a fundamental pre-condition. Figure 5 shows the range of the absolute difference for all possible wavelength combinations computed by stacking 540 matrices from the according number of spectra. Areas of very low variability (range < 0.05) are indicated by the black boxes. They can be detected likewise in lab and field spectra. On the whole, variability is higher for the visual and near infrared part of the spectrum in lab spectra. This is due to the artificially induced moisture variability. In nature, oven-dried soils do not exist whereas a moisture content of 13% at surface level is the opposite extreme which will only be encountered shortly after heavy rainfall or in relief depressions. In Table 2, centre wavelengths for stable areas for lab and field spectra, their mean absolute difference and standard deviations are compiled. Mean values are generally very low. Lab and field mean values differ on average by 0.0027, where mean values of field spectra are smaller.

To assess the effect of spectral resolution on stable features a resampling from 1 to 5, 10, 20 and 50nm at full width half maximum (FWHM) was performed. But resampling was not able to resolve stable areas in either case. Therefore, identified centre wavelengths (Table 2) will be referred to as initial points for synthetic endmember extraction from image date in the following.

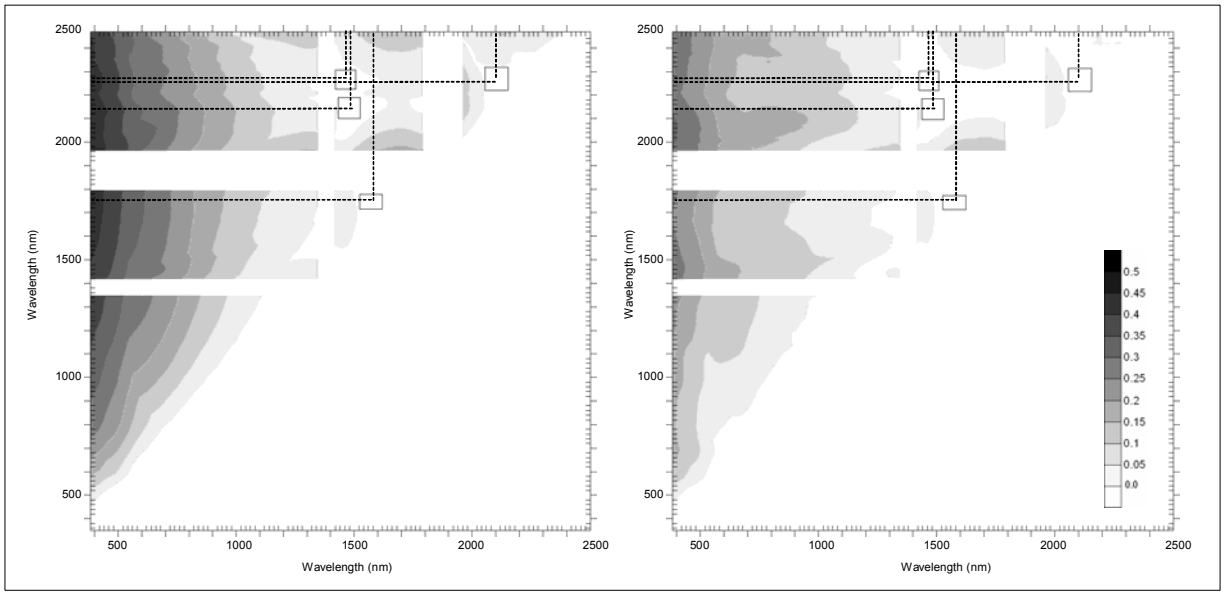

Figure 5: Range of absolute difference for all lab spectra (135 samples with 4 moisture contents and 3 zenith angles)

Table 2: Centre pixels for stable areas in lab and field spectra, their mean and standard deviation.

	Stable area wavelengths (nm)	Offset	Mean		St.dev.	
			Lab	Field	Lab	Field
1	1581 / 1750	169	0.009682	0.006820	0.002718	0.002713
2	1489 / 2151	662	0.015598	0.013637	0.007404	0.007828
3	1464 / 2243	779	0.014920	0.009147	0.009381	0.007169
4	2107 / 2264	157	0.011953	0.011937	0.007382	0.002946

The generation of look-up tables by linear regression analysis was done with a third of the entire data set. The sample of spectra used for the preparation of look-up tables was drawn arbitrarily from an adequate stratified sample (4 soil moisture classes). The remaining two-thirds of the data were subsequently used for validation purposes.

Results of the linear regression analysis for lab and field spectra are illustrated in Figure 6. The curves indicate R^2-values over the entire spectral range deploying all initial wavelengths points. In principle, R^2-values exceed 0.8 for lab spectra disregarding the spectral range from 350 to 450 nm which is acquired only by the ASD spectroradiometer. Very strong correlations exist between 1000 to 2500nm with $R^2 \geq 0.95$. Correlation for field spectra is generally lower, increasing from an R^2 of 0.7 at 450nm to an R^2 of 0.92 to 0.95 at 1000nm in VNIR range. R^2-values in SWIR I and II are about the same for lab and field spectra. Owing to autocorrelation R^2-values reach 1.00 at wavelengths of initial points.

In Figure 7 (a, b) two synthetic spectral endmembers are displayed along with the two corresponding spectra from HyMap image data acquired in August 2004 and May 2005. For the presented examples, pure HyMap pixels of bare soil were preferably selected. That way, differences between SY-SEM and HyMap spectra are expected to be rather low and may serve as an indication of method operation. On top of the graphs, R^2-values are marked with small crosses to point out the quality of linear regression for endmember generation. As expected from R^2-values, deviations between HyMap spectra and SY-SEM are largest in the VNIR region but spectra also diverge at 1900 to 2100nm. Since no field data was collected during data acquisition, it is difficult to judge the purity of the selected pixels as well as artefacts from atmospheric correction.

Thus, for validation purposes, synthetic spectral endmembers were also computed from lab spectra of the two-third of the data set which was not used for the generation of the look-up tables. As shown in Figure 7 (c, d), SY-SEM match very well over the entire curve with the input spectra. Figure 7a illustrates the spectral curves of a soil sample with 7.2M.-% water content. Figure 7b shows results for an oven dried soil respectively. These results lead us to the assumption, that synthetic spectral endmembers can be derived from image data relatively independent of inherent moisture conditions.

CONCLUSIONS AND OUTLOOK

The presented article suggested a new algorithm for deriving synthetic soil endmembers to support linear unmixing approaches. In contrast to common practise where endmember are obtained from either, lab spectra, field spectra or image data, we suggested a method which allows the derivation of synthetic endmembers for soil on basis of self-generated look-up tables and pre-classified image data on a pixel per pixel basis.

Most available methods for endmember extraction permit reasonable results but they are often very time-consuming. Generally, that fact has been addressed by optimizing the processing procedure e.g. with parallel processors (see introduction). In this article we suggest to reduce the time expenditure by an optimization of available input data that is previous knowledge of certain spectral features.

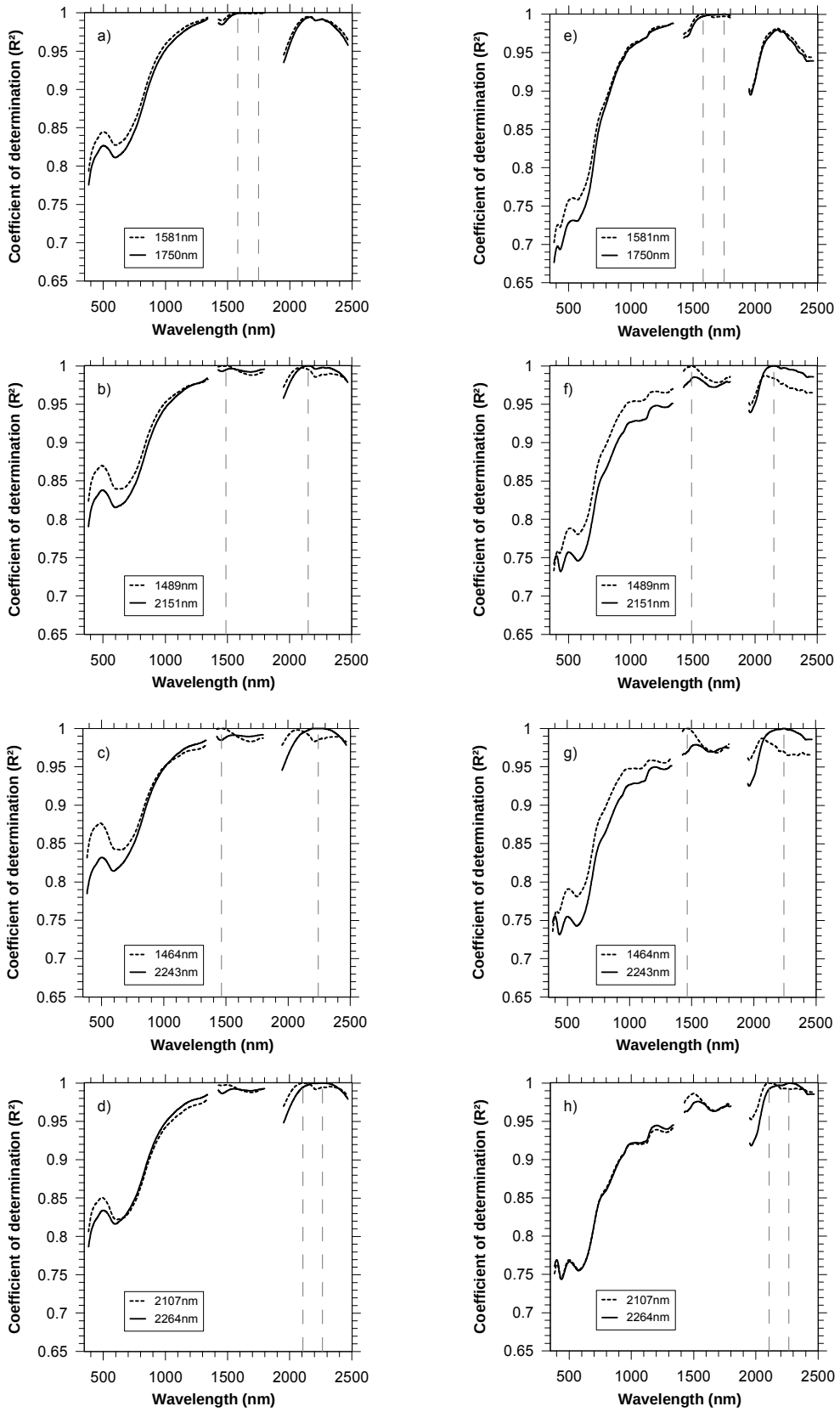

Figure 6: Coefficient of determination for reflectance values at initial points and all other reflectance values – lab spectra (a-d) and field spectra (e-h)

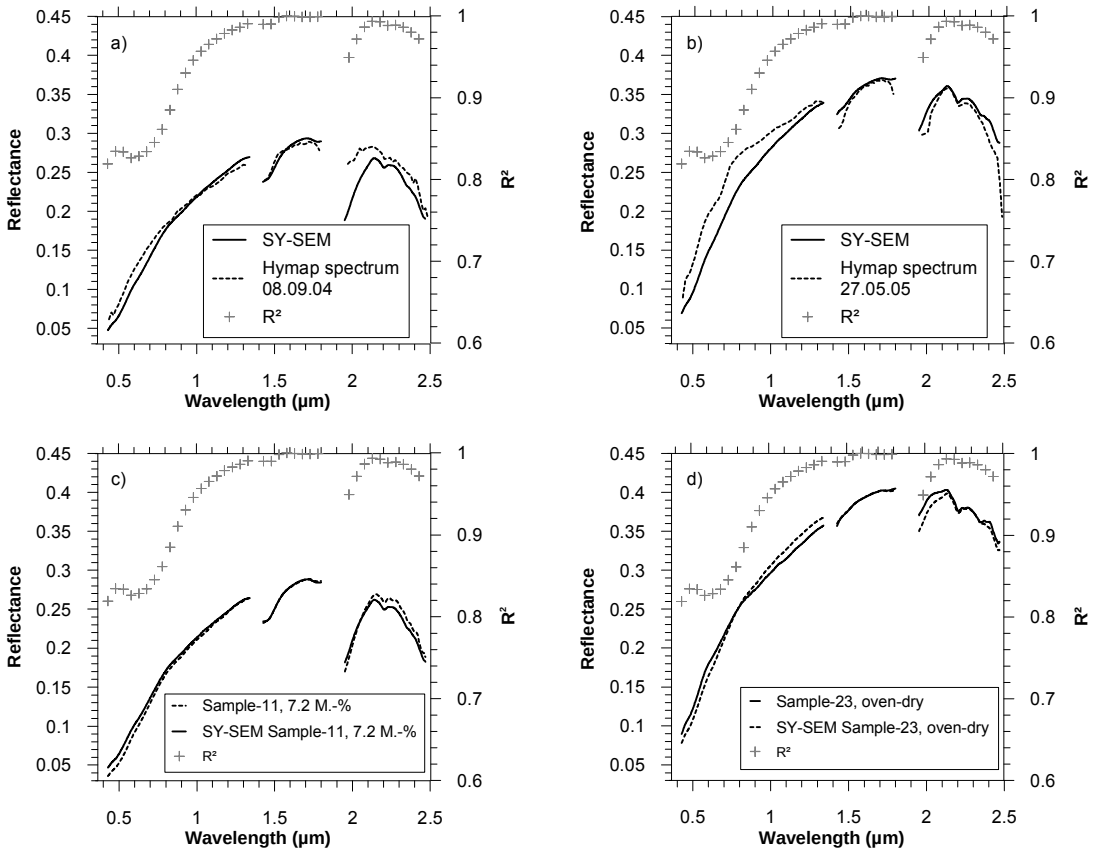

Figure 7 (a, b): Synthetic spectral endmember (SY-SEM) derived from HyMap image data and R²-values (cross) as an indicator of correlation accuracy.

Figure 7 (c, d): Synthetic spectral endmember (SY-SEM) derived from lab spectra of the validation data set and R²-values (cross) as an indicator of correlation accuracy.

Therefore, we developed a procedure to identify stable spectral features on the basis of lab and field spectra. A systematic analysis of spectra from 135 soil samples collected at DEMMIN test site and measured at varying moisture rates and illumination angles in the lab has been executed. Results of the study reveal wavelengths pairs of very low variability within the spectrum if the difference between a pair of wavelengths is calculated. Further, we showed that these wavelengths are strongly related with absolute reflectance values in SWIR I and SWIR II. R²-values from lab spectra are above 0.8 in the VNIR region, still explaining 80% of the variability within the data set. Based on these findings we set up a procedure for the synthesis of endmembers with quality measures for all modelled spectra.

Since data was derived from DEMMIN test site the approach is at this time limited to similar environments in Europe. However, its operation will be evaluated with soil spectra from other environments worldwide to test its transferability. An assessment of synthetic endmembers in linear unmixing and comparison with conventional approaches is currently in process.

These results form the basis for the extraction of soil input parameters such as clay from hyperspectral image data for erosion modelling.

ACKNOWLEDGEMENTS

The authors wish to thank Sören Haubrock and Sabine Chabrillat from the GeoForschungsZentrum in Potsdam for offering the opportunity to do the reflectance measurements in their spectros-

copy lab. We also thank Ms. Dagmar Schulz and her colleagues from the Central Laboratory of ZALF who did the physical and chemical analysis of our soil reference data base.

REFERENCES

i Adams J B, Smith M O & P E Johnson, 1986 Spectral mixture modeling – A new analysis of rock and soil types at the Viking Lander 1 site. Journal of Geophysical Research, 91: 8098-8112

ii Adams J B, Smith M O & A R Gillespie, 1993 Imaging spectroscopy: Interpretation based on spectral mixture analysis. In: Remote geochemical analysis: Elemental and mineralogical composition, edited by: C M Pieters & P A Englert (Cambridge: Cambridge University Press), 145-166

iii Boardman J W, Kruse F A & R O Green, 1995. Mapping target signatures via partial unmixing of AVIRIS data. In: Summaries of JPL Airborne Earth Science Workshop, Pasadena, CA.

iv Haubrock S, Chabrillat S & H Kaufmann, 2005. Application of hyperspectral imaging for the quantification of surface soil moisture in erosion monitoring and modelling. In: 4th EARSeL Workshop on Imaging spectroscopy. New quality in environmental studies. edited by B Zagajewski & M Sobczak (EARSeL, Warsaw), 197-206

v Hill J & B Schütt, 2000. Mapping complex patterns of erosion and stability in dry Mediterranean ecosystems. Remote Sensing of Environment, 74, 557-569

vi Hurtig Th, Fukarek F & J Stübs, 1957. Physische Geographie von Mecklenburg. (VEB Deutscher Verlag der Wissenschaften) pp .252

vii Ichoku C & A Karnieli, 1996. A review of mixture modelling techniques for sub-pixel land cover estimation. Remote Sensing reviews., 13, 161-186

viii Gruninger J, A. J. Ratkowski and M. L. Hoke. "The Sequential Maximum Angle Convex Cone (SMACC) Endmember Model". Proceedings SPIE, Algorithms for Multispectral and Hyperspectral and Ultraspectral Imagery, Vol. 5425-1, Orlando FL, April, 2004.

ix Moreno J, Menenti M & R Richter (2001) Assessments of inputs to land surface process models derived from hyperspectral multi-angular data. In: Proceedings of DAISEX Final Results Workshop, ESTEC, 15-16 March, ESA SP-499

x Morgan, R P C, J N Quinton & R J Rickson 1992. EUROSEM documentation manual (Version 1). (Silsoe College), Silsoe

xi Morgan, R P C, J N Quinton & R J Rickson 1993. EUROSEM – a user guide (Version 2). (Silsoe College), Silsoe

xii Neville R A, Staenz K, Szeredi T, Lefebvre J & P Hauff, 1999. Automatic endmember extraction from hyperspectral data for mineral exploration. In: 21st Can. Symp. Remote Sensing, Jun. 1999, 21–24.

xiii Plaza A, Martínez P, Pérez R & J Plaza, 2004. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. IEEE Trans. Geosci. Remote Sens., vol. 42, no.3, 650–663.

xiv Schmidt J, 1996. Entwicklung und Anwendung eines physikalisch begründeten Simulationsmodells für die Erosion geneigter landwirtschaftlicher Nutzflächen. Berliner Geographische Abhandlungen, Heft 61, Im Selbstverlag des Instituts für Geographische Wissenschaften.

xv Von Werner, M, 1995. GIS-orientierte Methoden der digitalen Reliefanalyse zur Modellierung von Bodenerosion in kleinen Einzugsgebieten. Dissertation, Freie Universität Berlin

xvi Winter M E, 1999, N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data.- In: Proc. SPIE Conf. Imaging Spectrometry V, vol. 3753: 266–277.