HYSOMA: AN EASY-TO-USE SOFTWARE INTERFACE FOR SOIL MAPPING APPLICATIONS OF HYPERSPECTRAL IMAGERY

S. Chabrillat [a,] *, A. Eisele [a], S. Guillaso [b], C. Rogaß [a], E. Ben-Dor [c], H. Kaufmann [a]

[a] Section 1.4 Remote sensing, GFZ German Research Center for Geosciences, Telegrafenberg, 14473 Potsdam, Germany – (chabri, eisele, rogass, charly)@gfz-potsdam.de

[b] Computer Vision and Remote Sensing, Berlin University of Technology, Franklinstr. 28/29, 10587 Berlin, Germany – stephane.guillaso@tu-berlin.de

[c] The Department of Geography and the Human Environment, Tel Aviv University, Ramat Aviv, PO Box 39040, Tel Aviv 69978, Israel – bendor@post.tau.ac.il

KEYWORDS: Hyperspectral, soil mapping, automatic algorithms.

ABSTRACT:

In the frame of the EU-FP7 EUFAR (European Facility for Airborne Research) project, higher performing soil algorithms are being developed as demonstrators for end-to-end processing chains with harmonized quality measures. For this, a review of existing soil algorithms and methodologies based on soil spectroscopy currently successful for the identification and prediction of soil properties was performed. Based on this review, the HYSOMA (Hyperspectral SOil MApper) interface was developed at the GFZ German Research Centre for Geosciences in the Remote Sensing section. It is a software package written in the IDL language for fully automatic generation of semi-quantitative maps of surface soil moisture content, soil organic carbon (SOC) content, and soil mineralogical content (e.g. iron oxide, clay mineral , carbonate).
The main motivation for HYSOMA development is to provide non-expert users with a suite of tools that can be used for soil applications. Also, HYSOMA tries to incorporate modern hyperspectral algorithms with an easy-to-use graphical interface based on simple menu-driven functions. It is achieved by providing a software with basic image file import based on ENVI format, soil mask option removing water and vegetation pixels based on simple water and vegetation indexes, and performing soil functions based on analytical and empirical algorithms where no user input data (e.g. spectral libraries, ground truth data) is needed. Additional tools that allow fully quantitative soil mapping are implemented for more experienced users. In this paper, we present the development strategy of the HYSOMA toolbox, current status, and examples of derived soil maps.

1. INTRODUCTION

Can we develop automatic algorithms for soil mapping based on spectral reflectance? This question comes nowadays at a high relevance level with the upcoming launch of the next generation of hyperspectral satellites such as the EnMap (Germany), Hyper-X (Japan), HyspIRI (US), HypXIM (France) and the increasing demand for the availability/accessibility of hyperspectral soil products from non-expert hyperspectral users. Indeed, although the use of hyperspectral imagery has had a very high increase in the past 10 to 20 years, it is a common assumption that hyperspectral analyses require an experienced user, as well as expensive specialized software packages. The lack of availability of automatic easy-to-use toolboxes, accessible for non-experts, has been a limitation to the further development of the use of hyperspectral imagery. As a result, automatic toolboxes are currently in development in Germany, in Israel, in Australia for the quantification of chemical and physical soil properties based on spectral reflectance. The availability of such automatic or semi-automatic toolboxes would allow non-expert users to access soil hyperspectral products to be used as input for e.g. further Earth's surface processes monitoring, soil resources evaluation, erosion and land degradation modelling, hydrological cycle modelling.

EUFAR (European Facility for Airborne Research) is an integrating activity funded by the European Commission under FP7. EUFAR works to coordinate the operation of instrumented aircraft and hyperspectral imaging sensors, exploiting the skills of experts in airborne measurements in the fields of environmental and geo-sciences, in order to provide researchers with the infrastructure best suited to their needs. A Joint Research Activity in EUFAR is the HYQUAPRO (Quality layers for airborne hyperspectral imagery and data products) project that aims at developing quality indicators and quality layers for airborne hyperspectral imagery for higher level data products. In HYQUAPRO higher performing soil algorithms are developed as demonstrators for end-to-end processing chains with harmonized quality measures. The higher performing algorithms must be tested and validated, and subsequently be integrated in the DLR (German Aerospace Center) Processing and Archiving Facility.

* Corresponding author.

2. REVIEW OF SOIL ALGORITHMS AND RECOMMENDATIONS FOR IMPLEMENTATION

2.1 Review of soil algorithms

At first a literature review of existing and current state-of-the-art methodologies used for soil parameter extraction based on imaging spectroscopy was performed. The algorithms were summarized and ordered in terms of development status and pre-knowledge needed (Table 1).

Algorithm	Output	Status	Pre-knowledge needed
BRDF-Radiative Transfer modelling	Soil-leaving reflectance	n/a	Soil geometry, illumination, albedo
Electronic and vibrational transitions theory	Calculation of absorption features	n/a	Molecule structure, replacement ion, energy fields
Spectral Mixture Analyses	Abundance image fraction. Quantification of soil type	r	End-member selection or spectral library
Spectral Feature Fitting	Mineral identification	r	Spectral library
Modified Gaussian Modelling	Quantification of soil biophysical properties	r-fdn	-
ENVI classification techniques	Mineral identification (SAM) or quantification (MF)	r	End-member selection
Spectral Indices	Semi-quantification of soil minerals or soil biophysical properties	r	-
Statistical Regression Analysis (LR, PLS)	Quantification of soil type, soil minerals, soil biophysical properties	r	In-situ data
Artificial Neural Network	Quantification of soil type	r-fdn	Learning/training vectors

Table 1. Review of existing and state-of-the-art algorithms used in hyperspectral imagery for soil applications: n/a, non available; r, robust; fdn, further development needed (source: EUFAR Deliverable DJ2.4.1, 2010).

2.2 Development of the HYSOMA interface: HYperspectral SOil MApper

The development work is based on the demand to provide a solution, balanced between the requirements for higher performing soil algorithms and the fact that the software shall be mainly addressed to non-expert users. Thus, HYSOMA software is focused on providing a suite of tools to users that are automatic methodologies using multiple algorithms to identify key soil parameters such as soil moisture, soil organic carbon, as well as different soil minerals (iron oxides, clay minerals, and carbonates), where no user input data (e.g. spectral libraries, ground truth data) is needed. These automatic methodologies that do not need input data can be classified into three groups: Spectral band analyses (continuum modelling, absorption area/depth, spectral slopes, etc..), spectral indices calculations, and Gaussian modelling methods. The drawback of these methodologies is that they provide at the best a semi-quantitative estimation of the soil parameter studied. Unfortunately, most quantitative methods like e.g. empirical approaches Partial Least Square (PLS) and Support Vector Machine (SVM) or spectral unmixing methods that are very widely used in soil sciences application, are user-data set driven or need a spectral library. Thus, for a more accurate determination of soil properties and more quantitative soil mapping possibilities, an additional "user custom" option that allows incorporation of e.g. users own models was implemented.

The source of the HYSOMA code is developed as standalone IDL software to allow easy implementation in the DLR PAF. Other reasons that support the choice of the IDL language are, a) within the hyperspectral community, it is very widely used since the most commonly used hyperspectral image processing software ENVI (www.itt.com) is programmed in IDL; b) outside the hyperspectral community for non-expert users that do not have an ENVI license, such a software can be executed as a binary version using the free IDL virtual machine under various operating systems including Linux, UNIX, Mac OS X and Windows. The HYSOMA code includes block programmation for handling of large data files, and multi-processor capabilities for fast calculations.

3. HYSOMA STRUCTURE AND METHODS

3.1 Conceptual framework

The HYSOMA processing flow is presented schematically in Figure 2. In a first step (Preprocessing) HYSOMA reduces the spectral range to the number of bands, which are suitable for any of the further analyses. Bands which do not provide meaningful values are excluded from the available band list.

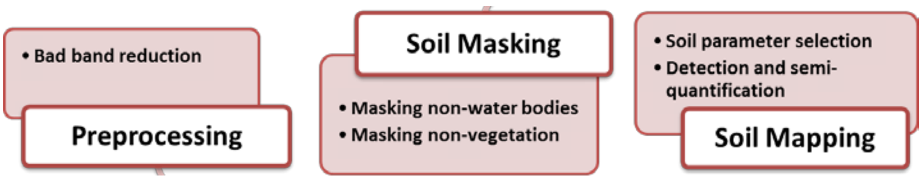

Figure 2. The HYSOMA (Hyperspectral Soil Mapper) conceptual framework.

In a second step (Soil Selection) , also called mask, the software selects spectrally soil dominated pixels and eliminates from the image all pixels which are not dominated by a soil signature. This is realized in HYSOMA by masking and excluding water pixels and vegetation pixels, in both vital and dry condition. Water dominated pixels are excluded through the Normalised Difference Red Blue Index NDRBI as suggested by Carter (1991) and Zakaluk and Ranjan (2008), which simply uses the ratio of the difference and the sum between the red (660 nm by default) and blue band (460 nm by default). Often specific local water characteristics or sensor spectral range limitations demand wavelength settings different from the defaults to detect water bodies. Thus, the ratio associated wavelength can be chosen manually in the HYSOMA-interface, such as using an SWIR band instead of the red band in the ratio. To mask the remaining non-soil pixels, HYSOMA identifies vegetation dominated regions and excludes them from further processing. This is accomplished by using two vegetation indices which are broadly used in remote sensing, the Normalised Difference Vegetation Index NDVI (Tucker, 1979) to mask green photosynthetic vegetation, and the normalised Cellulose Absorption Index nCAI (Nagler et al., 2003) to mask non-photosynthetic activity.

Finally, HYSOMA Soil Mapping module performs soil functions, and produces soil maps based on the spectrally soil dominant pixels, which remain from the soil masking procedure. The one-dimensional grey value maps can easily be imported and visualized into any image processing software for visualisation. In total, for the 5 soil selected parameter, 11 algorithms are proposed (summarized in table 2), and 11 map files are created. Additionally, map files associated with the soil selection procedure (water map, NDVI map, CAI map, soil dominant pixels map) are saved. Also, an HYSOMA run report file is saved in the output directory.

3.2 Graphical user interface

The main purpose of HYSOMA is to ease as much as possible the generation of soil maps offering multiple choice of algorithms for every user and simple inputs-outputs procedure, targeted for non-experts hyperspectral users. Therefore, HYSOMA is based on easy-to-use graphical user interfaces that are very simple and basically provide only selectors for input image file and output directory. The main graphical user interface of HYSOMA is depicted in Figure 3.

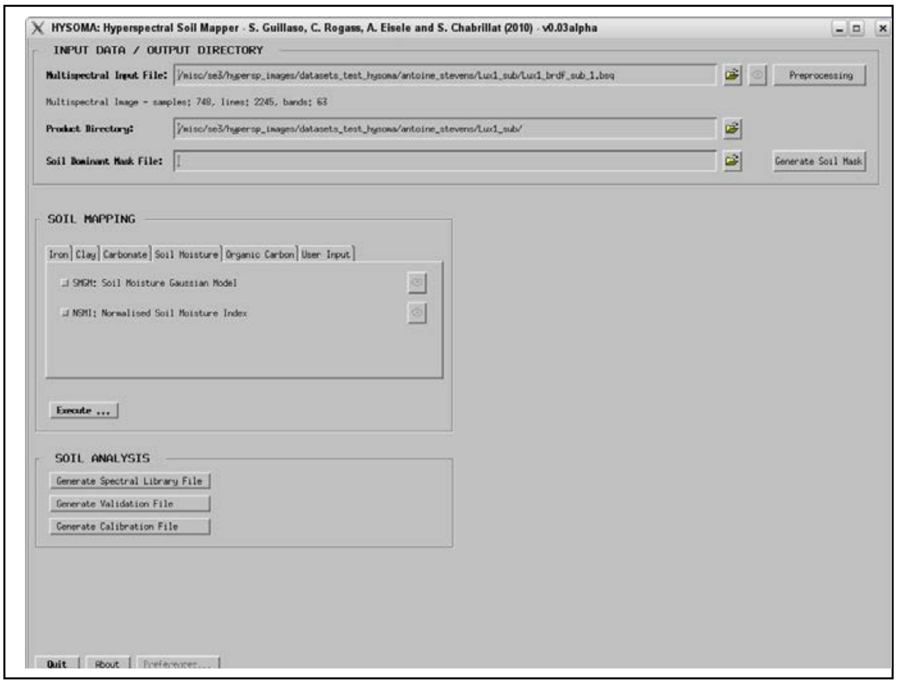

Figure 3. HYSOMA main graphical user interface.

The HYSOMA reads input image files in the ENVI format (image file and associated header) and standard HDF5 format. After finishing the data processing, HYSOMA automatically generates soil maps files in the given output directory.

3.3 Soil mapping tools

Soil functions for identification and semi-quantification: The selected soil algorithms which were integrated into HYSOMA are presented in table 4 with their associated soil chromophores and references. To derive spectral feature absorption depths, HYSOMA applies an automatic spectral feature analyses technique that fits a continuum between two defaults shoulders that are within a spectral range well beyond the size of the absorption feature to minimize calculations errors. The methods SOC1, 2 and 3 are defined in Bartholomeus et al., 2008, and are self-defined indices based on spectral area and spectral slope calculations.

Soil chromophores	Soil algorithm	Spectral Region (nm)	Estimated soil parameters
Clay Minerals Al-OH content	Clay index (SWIR FI)	2209, 2133, 2225	Clay mineral content (Levin et al., 2007)
	Clay absorption depth	2150 - 2250	Clay mineral content
Iron Oxides Fe_2O_3 content	Iron index (RI)	477, 556, 693	Hematite content (Madeira et al., 1997 ; Matthieu et al., 1998)
	Iron absorption depth 1	450 - 630	Iron oxide content
	Iron absorption depth 2	750 - 1040	Iron oxide content
Carbonates Mg-OH content	Carbonate absorption depth	2300 - 2400	Carbonate content
Soil Moisture Content	Moisture index (NSMI)	1800, 2119	Soil moisture content (Haubrock et al. 2008a, 2008b)
	Gaussian modelling (SMGM)	~1500-2500	Soil moisture content (Whiting et al. 2004)
Soil Organic Carbon	Band analysis SOC 1	400 - 700	Organic matter content (Bartholomeus et al., 2008)
	Band analysis SOC 2	400 , 600	Organic matter content (Bartholomeus et al., 2008)
	Band analysis SOC 3	2138 - 2209	Indirect organic matter content (Bartholomeus et al., 2008)

Table 4. Soil algorithms available for fully automatic soil mapping in HYSOMA, grouped per soil chromophore.

Additional user custom option: For individual feature analysis requirements from experimented users an additional absorption depth option was included into the HYSOMA software, which allows user defined spectral feature selection via an editable left and right shoulder tab for a continuum removal band depth creation. Additionally, in the user input option is proposed fully quantitative soil mapping based on input regression file where users can apply directly on the image their own regression models such as PLS prediction equation. Input regression files are in csv format.

3.4 Soil analyses tools

Spectral library: This option allows experimented or non-experimented users to extract individual spectra from the input image based on their geographic coordinates, and to put them in a spectral library file. Also, HYSOMA reads as input data file not only hyperspectral images but spectral library file in ENVI format, on which all soil functions can be then performed.

Calibration file: This option allows experimented users to perform fully quantitative mapping using input field data for calibration. This input field data option allows calibrating automatically generated soil semi-quantified maps (output from one of the algorithm in table 2) with field measurements. Two methods are proposed. Either the users enter directly a field measurement file with name of field location, coordinates X,Y and absolute value of soil parameter, and HYSOMA performs the calibration and delivers as output a quantitative soil map file, or the users give as input already calculated gains and offsets for calibration.

Validation file: This option allows the users to extract from HYSOMA output soil maps the soil parameter values of individual points based on their geographic coordinates.

4. EXPERIMENTAL RESULTS

4.1 Soil organic carbon mapping

HYSOMA has been tested for different data sets from different sensors. Soil organic carbon mapping was tested based on a Airborne Hyperspectral Spectrometer AHS (Fernández-Renau et al., 2005) hyperspectral dataset courtesy of A. Stevens, Université Catholique de Louvain over a test site in Grand-Duchy of Luxembourg. The AHS-160 from the Remote Sensing Laboratory at INTA (www.inta.es), Madrid, Spain, provided spectral images with 63 spectral bands from ~450 to 2500 nm. Due to a reduced signal-to-noise ratio in the SWIR we focused our analyses on the VNIR part of the spectrum distributed on 20 contiguous spectral bands from 442 to 1019 nm with a spectral resolution between 27 and 30 nm. The spatial resolution is 2.6 x 2.6 m (Stevens et al., 2010). SOC 1 uses the inverse of the area of the absorption feature between 400 and 700 nm, and expresses the amount of absorption. SOC2 uses the inverse slope of the reflectance between 400 and 600 nm. HYSOMA showed coherent results for both SOC 1 and SOC2 methods (Figure 5) and high-medium correlation with the organic carbon content from 81 ground truth samples (EUFAR deliverable D.J.2.4.2). SOC 3 which covers the SWIR region of the spectrum cannot be used with this dataset.

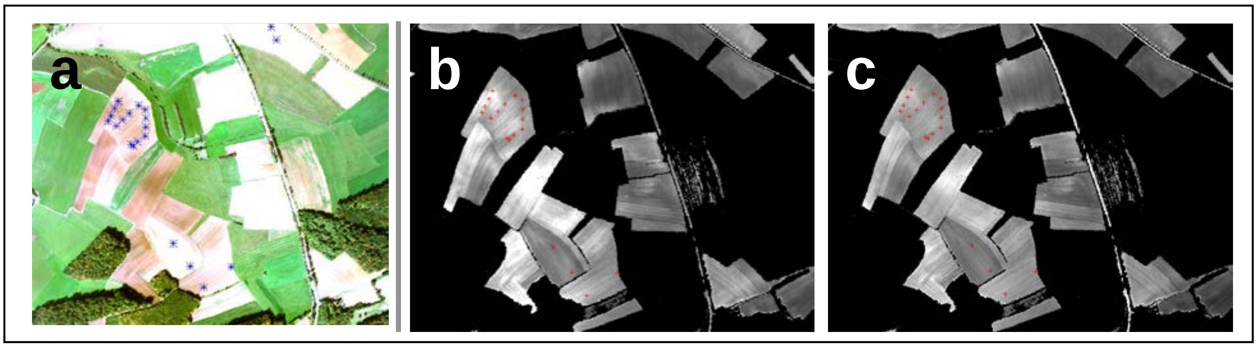

Figure 5. Test site in Grand-Duchy of Luxembourg: a) True color RGB of AHS-160 images (subset); b) SOC1 soil map; c) SOC2 soil map. Grey-color coded ranges from white (high SOC content) to black (low content). The stars locate field ground truth locations.

4.2 Mineralogical mapping

Iron mapping was tested based on a HyMap hyperspectral image from a test site in South Eastern Spain (Cabo de Gata) recorded in June 2005 (Chabrillat et al., 2005). The HyMAp sensor provides 126 spectral bands across the 450-2500 nm region with contiguous spectral coverage and bandwidths between 12 and 17 nm (Cocks et al., 1998). Figure 6 shows the results for two methods of iron oxide determination. The first method measures absorption depth of the iron absorption feature centred around 525 nm, and the second method calculates a iron spectral index based on the TM Redness Index defined by Madeira et al. (1997) and adapted for hyperspectral wavelengths channels. The Redness Index is known to account mainly for hematite content. Figure 6 shows that both soil functions demonstrate coherent results with the geological knowledge of the area and mineralogical information from the iron abundance map derived from HyMap imagery and field calibration (Richter et al., 2007).

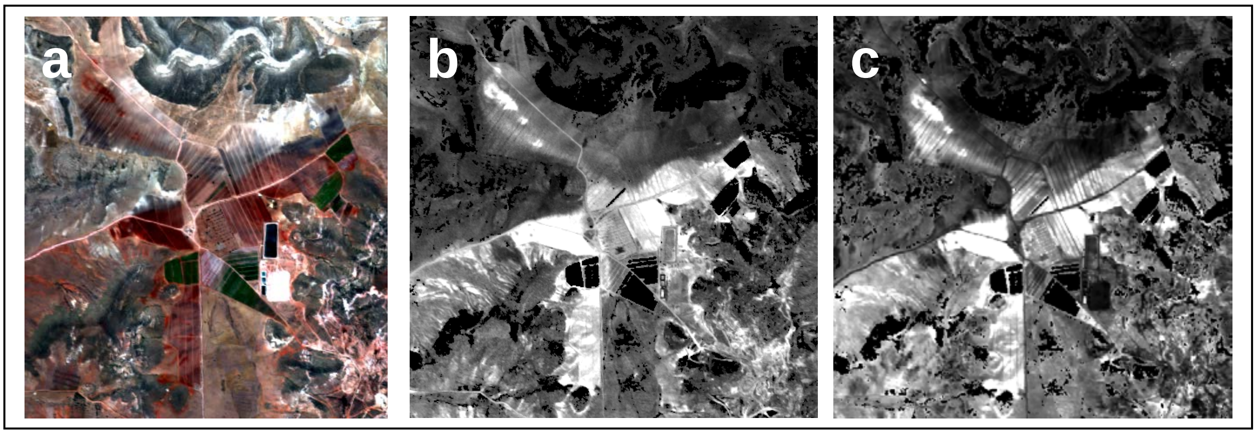

Figure 6. Test site in Cabo de Gata National Park, Spain: a) True color RGB of HyMap images (subset); b) iron oxide soil map based on the iron spectral feature analysis (absorption depth between 460 and 620 nm); c) hematite soil map based on the Redness Index (Madeira et al., 1997). Grey-color coded ranges from white (high iron content) to black (low content).

5. AVAILABILITY / TERMS OF USE

HYSOMA represents a development that is made at GFZ within the EUFAR project and the current prototype is in the last phase of test and development. Part of the EUFAR contract, HYSOMA will be integrated in the DLR PAF. HYSOMA is an experimental platform for soil applications of hyperspectral imagery that has got some limitations such as e.g. it reads only ENVI and HDF5 standard file, it includes right now no image visualisation, and the processing methods are fixed. We will probably add further functionalities in the future.

HYSOMA has been entirely programmed in IDL (Interactive Data Language) and will be distributed in a pre-compiled version (sav file). To run it, an IDL license or the freely available IDL virtual machine is necessary (www.ittvis.com/idlvm). HYSOMA will be provided free-of-charge via the internet on the GFZ remote sensing section web site. The HYSOMA interface will be also included into the EUFAR toolbox in 2012. The usage will be completely free for non-commercial and educational purposes. Detailed license regulations will be available when HYSOMA is released.

6. SUMMARY AND OUTLOOK

After an intensive literature review and following expert suggestions from the EUFAR Expert Working Group Workshop on hyperspectral soil applications held at Potsdam, Germany, 15-16 April 2010, the GFZ has undertaken the development of higher performing soil algorithms with three criteria: use of methodologies where automation is possible, choice of multiple algorithms by the users, implementation of a user custom option to incorporate user-driven applications and more quantitative mapping. Thus, the HYperspectral SOil MApper (HYSOMA) interface was developed in the IDL language. It is an experimental platform for soil mapping applications of hyperspectral imagery. The main motivation for HYSOMA development is to provide non-expert users with a suite of tools that can be used for soil applications.

The software focuses on the identification and semi-quantification of key soil parameters such as soil moisture, soil organic carbon, as well as different soil minerals (iron oxides, clay minerals, and carbonates), where no user input data (e.g. spectral libraries, ground truth data) is needed. In total, for the five soil parameters selected, 11 algorithms for soil mapping are proposed. Additionally to the HYSOMA soil functions focused on the identified five key soil parameters, a user custom option was added, that allows to perform user-driven spectral band calculations to identify some additional user-driven soil parameters. This additional function allows also fully quantitative mapping where the user can apply on the image his own regression models (such as PLS prediction equation), or input field measurements for calibration.

As a conclusion, the development of higher performing soil algorithms for non-expert users based on hyperspectral imagery lead to the challenging development of the HYSOMA interface. Content of the HYSOMA interface and resulting soil products validated with different data sets demonstrated the accuracy of the algorithms. HYSOMA has several disadvantages as it could not include common algorithms used for soil science mapping such as PLS and SVM methods that would need user-driven expert knowledge and a heavier user interface. Also, it can be a bit slow depending on the image size. Nevertheless, HYSOMA provides fully automatic generation of semi-quantitative maps of soil mineralogy and moisture content, organic carbon content, which at this date is a big and necessary step toward the implementation and future development of soil toolboxes. These toolboxes next generation will allow easy access to L3 geoscience products as hyperspectral imagery will become more and more used by 'non-expert' hyperspectral users as its accessibility will soon increase exponentially with the upcoming launch of the next hyperspectral satellite in 2015.

REFERENCES

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., and Spaargaren, O.S.P., 2008. Spectral reflectance based indices for soil organic carbon quantification. *Geoderma*, 145, pp. 28-36.

Carter, G.A., 1991. Primary and Secondary Effects of Water Content on the Spectral Reflectance of Leaves. *American Journal of Botany*, 78(7), pp. 916-924.

Chabrillat, S., Kaufmann, H., Escribano, P., and Palacios-Orueta, A., 2005. Land degradation monitoring: Spectral variability in a semi-arid Mediterranean ecosystem (Natural Park Cabo de Gata-Níjar, Spain). In: *Proceedings of the 4th Workshop on Imaging Spectroscopy: New quality in Environmental Studies, Warsaw, Poland, 27-29 April 2005*, B. Zagajewski and M. Sobczak (eds), EARSeL & Warsaw University, Warsaw 2005, pp. 139-146.

Cocks, T., Jenssen, R., Steward, A., Wilson, I. and Shields, T., 1998. The HyMap™ airborne hyperspectral sensor: The system, calibration and performance. In: *Proceedings of the 1st European Association of Remote Sensing Laboratories (EARSeL) Workshop on Imaging Spectroscopy*, Zurich, Switzerland, 6-8 October 1998, EARSeL, Paris, pp. 37-42.

Fernández-Renau, A., Gómez, J. A., and De Miguel, E., 2005. The INTA-AHS system. In: *SPIE proceeding of Sensors, Systems and Next-Generation Satellites IX*, vol. 5978, pp. 471–478.

Haubrock, S.-N., Chabrillat, S., Kuhnert, M., Hostert, P. and Kaufmann, H., 2008b. Surface soil moisture quantification and validation based on hyperspectral data and field measurements. *Journal of Applied Remote Sensing*, Vol. 2, 023552.

Haubrock, S.-N., Chabrillat, S., Lemmnitz, C. and Kaufmann, H., 2008a. Surface soil moisture quantification models from reflectance data under field conditions. *Int. J. Remote Sensing*, 29(1), pp. 3-29.

Levin, N., Kidron, G.J. and Ben-Dor, E., 2007. Surface properties of stabilizing coastal dunes: combining spectral and field analyses. *Sedimentology*, 54, pp. 771-788.

Madeira, J., Bedidi, A., Cervelle, B., Pouget, M. and Flay, N., 1997. Visible spectrometric indices of hematite (Hm) and goethite (Gt) content in lateritic soils: the application of a Thematic Mapper (TM) image for soil-mapping in Brasilia, Brazil. *Int. J. Remote Sens.*, 18(13), pp. 2835-2852.

Mathieu, R., Pouget, M., Cervelle, B. and Escadafal, R., 1998. Relationships between satellitebased radiometric indices simulated using laboratory reflectancedata and typic soil color of an arid environment. *Remote Sens. Environ.*, 66, pp. 17-28.

Nagler, P.L., Inoue, Y., Glenn, E.P., Russ, A.L., and Daughtry, C.S.T., 2003. Cellulose absorption index (CAI) to quantify mixes soil-plant litter scenes. *Remote Sens. Environ.*, 87, pp. 310-325.

Richter, N., Chabrillat, S. and Kaufmann, H., 2007. Enhanced quantification of soil variables linked with soil degradation using imaging spectroscopy. In: *Proceedings of the 5th EARSeL Workshop "Imaging Spectroscopy: Innovation in Environmental Research"*, Bruges, Belgium, 23-25 April 2007, I. Reusen and J. Cools (eds), On CD-ROM, 7 pp.

Stevens, A., Udelhoven, T., Denis, A.,Tychon, B., Lioy, R., Hoffmann, L. And van Wesemael, B., 2010. Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. *Geoderma*, 158 (1-2), pp. 32-45.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.*, 8, pp. 127-150.

Whiting, M.L., Li, L., and Ustin, S.L., 2004. Predicting water content using Gaussian model on soil spectra, *Remote Sens. Environ.*, 89, pp. 535-552.

Zakaluk, R., and Ranjan, R.S., 2008. Predicting the leaf water potential of potato plants using R G B reflectance. *Canadian Biosystems Engineering*, 50(22).

ACKNOWLEDGMENTS

This work is funded by the European Commission under FP7 'Integrating Activity' EUFAR (www.eufar.net). The authors acknowledge the EUFAR-JRA2 HYQUAPRO partners for intensive discussions regarding the development of HYSOMA. We thank internal funding at the GFZ Potsdam that further promoted this work. The 1st author thanks Antoine Stevens and Sören Haubrock for sharing of Luxembourg and Welzow imagery and ground truth data for HYSOMA testing. Mike Whiting is furthermore gratefully acknowledged for the checking of our version of the SMGM algorithm and further discussions on HYSOMA.