TOPSOIL PROPERTIES ESTIMATION FOR AGRICULTURE FROM PRISMA: THE TEHRA PROJECT

R.Casa[1], R. Bruno[2], V. Falcioni[1], L. Marrone[1], S.Pascucci[3], S.Pignatti[3], S.Priori[1], F.Rossi[4], A.Tricomi[2], R.Guarini[5]

[1] Department of Agriculture and Forestry Sciences (DAFNE), University of Tuscia, Via San Camillo de Lellis, 01100 Viterbo (Italy)
[2] e-GEOS S.p.A., Via Tiburtina, 965 - 00156 Rome - Italy
[3] Institute of Methodologies for Environmental Analysis (IMAA)- Italian National Research Council (CNR), C. da S.Loja, 85050 Tito Scalo, Italy
[4] Dipartimento di Ingegneria Astronautica, Elettrica ed Energetica (DIAEE), University of Rome "La Sapienza", Via Eudossiana 18, 00184 Roma, Italy
[5] Italian Space Agency (ASI), Downstream & Application Services Department, Località Terlecchia snc – 75100 Matera, Italy

ABSTRACT

The project "Topsoil properties Estimation from Hyperspectral Remote sensing for Agriculture" (TEHRA), funded by the Italian Space Agency (ASI), aims at developing methods and algorithms for the estimation of soil properties of agronomic and environmental interest from PRISMA satellite hyperspectral data, that could support: 1) the adoption of more sustainable and climate-smart farming practices, e.g. through the implementation of precision agriculture applications; 2) monitoring in support of agricultural and environmental policies, e.g. related to climate change and for the encouragement of the adoption of practices preserving soil health.

In this paper, some results of the first year of the project are illustrated. They concern: 1) a scenario definition study; 2) studies on the confounding effect of soil moisture and crop residues; 3) exploitation of multi-temporal PRISMA data and 4) data fusion with proximal soil sensing.

Index Terms— Imaging spectrometry, soil mapping, hyperspectral, ASI

1. INTRODUCTION

The management of such complex systems as agroecosystems, can greatly benefit from the increased monitoring possibility offered by imaging spectroscopy of topsoil variables. Observation, monitoring and mapping of soils, a crucial element of the agroecosystem, is now increasingly possible by exploiting satellite imaging spectroscopy data. This offers the possibility to improve the agronomical management, and to support agricultural and environmental policy monitoring, primarily in the interest of environmental sustainability [1].

The increasing availability of high signal-to-noise ratio satellite imaging spectrometers on board on-going missions such as PRISMA [2] and EnMAP [3], as well as forthcoming ones, such as NASA-SBG (Surface Biology and Geology) [4] and CHIME (Copernicus Hyperspectral Imaging Mission for the Environment) [5], opens great perspectives for operational soil monitoring and mapping.

Building upon a considerable amount of research work carried out in the past two decades, illustrating the great potential of imaging spectroscopy in the context of regional to local soil properties retrieval [6], [7], the TEHRA project aims at targeting specifically key issues and constraints that currently limit the operational use of hyperspectral satellite data for supporting precision agriculture and environmental policies monitoring, through the development quantitative soil monitoring and mapping products generated by PRISMA.

The first year of the project, led by the University of Tuscia (Viterbo, Italy) and involving the National Research Council of Italy (CNR-IMAA) and the e-GEOS company (Italy), has included activities concerning: 1) a scenario definition study on the representativeness of surface topsoil for the cultivated rootzone; 2) studies on the confounding effect of soil moisture and crop residues on the retrieval of topsoil properties; 3) exploitation of multi-temporal PRISMA data and 4) data fusion of PRISMA with proximal soil sensing.

The main results of these activities are reported in the following sections.

2. SCENARIO DEFINITION STUDY

Spaceborne imaging spectroscopy of soils can only access information strictly limited to the soil surface, although the interest of the user communities would concern the topsoil rootzone, i.e. until a depth of 30-40 cm. It is often assumed that a homogeneous vertical distribution of soil properties occurs in the top rootzone layer, but the vertical variation of soil properties in cultivated soils would actually depend on the tillage system. A systematic review of these aspects is lacking in the current literature. Therefore, an analysis was carried out of data collected from studies in which the vertical variation of the main soil properties was reported, for four types of tillage systems: i) deep inversion ploughing (DP); ii) subsoiling (SUB); iii) minimum tillage (MT) and no-tillage (NT). Data were extracted from 43 papers with data sets suitable for the analysis of this study. NT showed the greatest vertical variation of Soil Organic Carbon (SOC), confirming the results of numerous papers and meta-analyses about the effects of tillage systems on soil organic matter. On the contrary, DP indicated a strong homogeneity in the top rootzone SOC, due to the inversion tillage and mixing of the crop residues. Linear models of SOC vertical variation were obtained for MT and SUB, whereas second order polynomial functions were calculated for DP and NT. The full cross-validation of the models provided R2 higher than 0.8 and errors (RMSE) from 1.55 to 3.23 g·kg-1, with similar efficiency for SOC and Total Nitrogen (TN). These models might be employed to extend the estimation of surface soil properties estimation from remote or proximal sensing to the whole topsoil layer of cultivated soil, provided a knowledge of the adopted tillage system is available [8].

3. CONFOUNDING EFFECT OF SOIL MOISTURE AND CROP RESIDUES

To tackle soil moisture confounding effect on the retrieval of other topsoil properties, a physically based estimation has been pursued in the TEHRA project. The MARMIT-2 model [9], [10] was used to generate simulated spectra of wet soils, by varying model parameters. These datasets were exploited to accomplish two objectives: (a) the estimation of soil moisture content, (b) the normalization of wet soils to dry spectra. Estimation of soil moisture was performed through machine learning regression, while various methodologies, including external parameter orthogonalization (EPO) [11] and autoencoder neural networks were explored to address normalization. Local soil spectral libraries (SSL), assembled by the project team, have been employed in order to develop and test the physically based soil moisture model MARMIT. The spectral libraries consisted of soil samples collected over a few sites in Italy (Maccarese, Pignola and Castelluccio) for which spectra were acquired in the lab by means of a spectroradiometer, during a natural soil drying process, concurrently measuring quantitatively gravimetric soil moisture. The Gradient Boosting Regressor trained on MARMIT generated databases was applied to PRISMA images to generate soil moisture maps (Figure 1). The generation of dry soil spectra from wet soil spectra was attempted by using both the EPO and autoencoder neural networks (Figure 2).

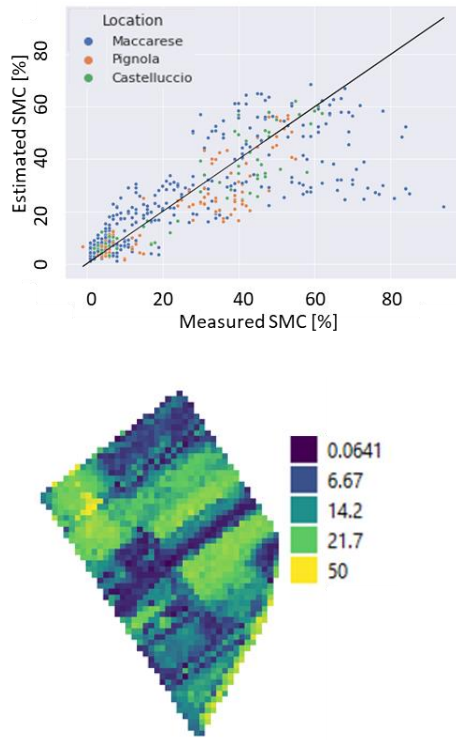

Figure 1. Retrieval of soil moisture content from PRISMA through Gradient Boosting Regressor (bottom) trained on MARMIT databases. Performances on Maccarese-Pignola-Castelluccio 2021 SSL (external SSL) is also reported (top).

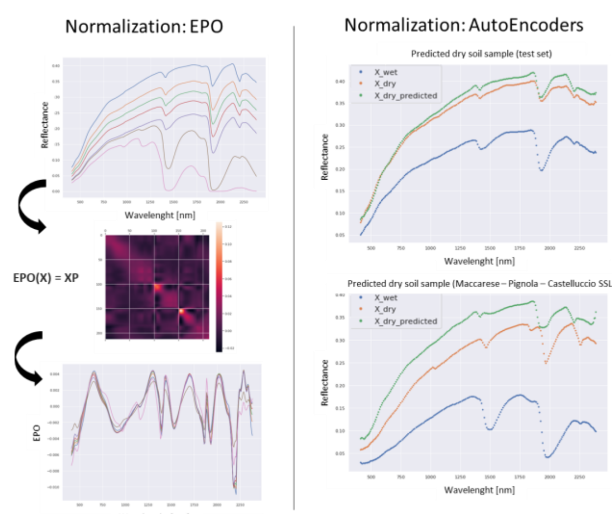

Figure 2. EPO [11] applied on dry and wet reflectance spectra generated by simulations (left). Autoencoder trained on simulations performed on MARMIT databases: sample from test set (top right) and sample (bottom right) from Italian test sites of Maccarese-Pignola-Castelluccio 2021 SSL.

The occurrence of crop residues, i.e. non-photosynthetic vegetation (NPV), such as stalks, stubble, and other senescent plant material) on the surface of agricultural soils also hinders the retrieval of topsoil properties. Spectral separation of the NPV and soil is challenging because NPV and soil show broad spectral similarity with few distinguishing features in the VNIR. Moreover, separation of the soil vs crop residues is highly dependent on soil moisture content and the age and decomposition state of crop residue, resulting in a low degree of accuracy when these conditions are unknown. NPV exhibits absorption features related to plant biochemical constituents, primarily lignin and cellulose (ligno-cellulose), in the shortwave infrared (SWIR) wavelengths, at 2100 and 2300 nm. In this context, PRISMA was used to identify crop residues by a linear regression model. The method exploits cellulose spectral features in the SWIR at 2096 nm characterizing it with continuum removal techniques (Figure 3).

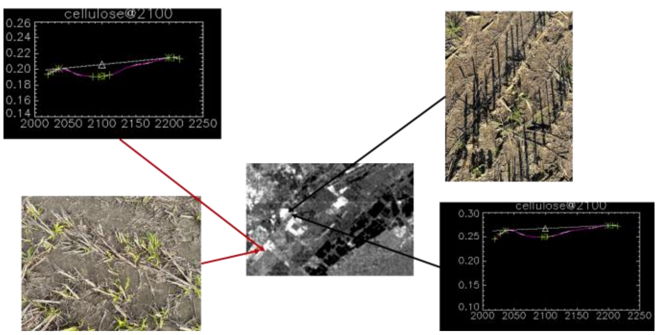

Figure 3. Continum Removal Analysis to assess cellulose absorption peak at 2096 nm.

4. EXPLOITATION OF MULTI-TEMPORAL PRISMA DATA

Capitalizing on the repeated satellite overpasses by spaceborne imaging spectrometers that are becoming available, such as PRISMA and EnMAP, there is great interest in exploiting multi-temporal observations to improve the prediction of soil properties by exploiting bare soil spectra collected in different soil moisture and roughness conditions. These are expected to provide more general information on the soil as compared to a single date. In the TEHRA project, repeated PRISMA bare soil reflectance data, acquired on some agricultural fields in the Jolanda di Savoia site in Northern Italy, were subject to the following computations: (a) the median spectra, (b) the spectra corresponding to the minimum value of the Soil Surface Moisture Index S2WI [12] and (c) to the maximum value of the Bare Soil Index [12]; (d) the spectra that correspond to the 90th quantile and (e) the 10th quantile of the reflectance for each band. These data were extracted following a pre-processing of PRISMA L2D data which included geometric co-registration correction based on Sentinel-2 images of close dates and spectra smoothing through the Savitzky-Golay filter and removal of bad bands, consisting of anomalous peaks and atmospheric features adsorption bands. The resulting spectra were used for the calibration of machine learning regression algorithms, such as Gaussian Process Regression (GPR) and support vector regression (SVR). The results showed that the minimum of the S2WI was the extraction method, and SVR was the algorithm, occurring most often in the best performing calibrated models. The property estimated with the best performance was clay, with a RMSE of 5.25% and a R2 of 0.89. SOC provided equally good results, with a RMSE of 0.84 g/kg and R2 of 0.86 [13].

5. DATA FUSION OF PRISMA WITH PROXIMAL SOIL SENSING

Another approach investigated in the TEHRA project deals with the coupling of hyperspectral remote data and proximal sensors, with data fusion techniques, to potentially increase the accuracy and spatial resolution of the soil properties prediction [7]. Two types of croplands have been used for this study: i) JOL- an arable field in Northern Italy (Jolanda di Savoia, Ferrara) of about 15 ha; ii) BRO- seven vineyards of a winery in central Italy (Brolio castle, Siena), for a total surface of about 30 ha. To reduce the dimensionality of the hyperspectral data and to preserve as much as possible their information content, a principal component analysis (PCA) of the spectra extrapolated from each image pixel was carried out. The first PCs (PC1) explained most of the variance of the images, therefore, it was selected for analysis. Regarding proximal electromagnetic induction, apparent electrical conductivity of shallower depth (ECa about 0-50 cm) has been used. Preliminary tests employed Regression Kriging, forward stepwise for $p < 0.5$, and Multiple Geographically Weighted Regression with gaussian weighting function, coupling a PRISMA PCA of the spectra with electrical conductivity of shallower depth (ECa1 about 0-50 cm) in an arable field and in vineyards. The results revealed the usefulness of including both bare soil and vegetation images, e.g. for the estimation of SOC and pH [14].

6. CONCLUSIONS

The project TEHRA has completed the activities planned for the first year. Work is ongoing on the different phases of the project concept plan, ranging from the scenario definition to the development of data driven machine learning and physically based modelling approaches for the estimation of topsoil properties from hyperspectral remote sensing. A good deal of relevant data has been assembled by the different teams of the consortium, both by carrying out ground field campaigns, by assembling soil spectral libraries and by acquiring a large number of PRISMA archive data over the TEHRA test sites in Central and Northern Italy. Activities are under way on the development of state of the art and innovative algorithms applied to the different retrieval methodologies. In the context of the international collaboration of the members of the THERA consortium, it is also planned to start exploring the use of EnMAP data as soon as they become available over the sites of interest.

7. ACKNOWLEDGEMENTS

This work was supported by the Italian Space Agency (ASI), by project TEHRA "Topsoil properties Estimation from Hyperspectral Remote sensing for Agriculture" [Contract/Agreement No. 2022-6-U.0, CUP 451 no. F83C22000160005]. Data generated by the TEHRA Consortium under a license from ASI Original PRISMA Product - © Italian Space Agency (ASI) - (2022).

11. REFERENCES

[1] J. Stafford, *Precision agriculture for sustainability*, 1st ed. Burleigh Dodds Science Publishing, 2018.

[2] S. Cogliati *et al.*, "The PRISMA imaging spectroscopy mission: overview and first performance analysis," *Remote Sens. Environ.*, vol. 262, no. April, 2021.

[3] L. Guanter *et al.*, "The EnMAP spaceborne imaging spectroscopy mission for earth observation," *Remote Sens.*, vol. 7, no. 7, pp. 8830–8857, 2015.

[4] E. N. Stavros *et al.*, "Designing an Observing System to Study the Surface Biology and Geology (SBG) of the Earth in the 2020s," *J. Geophys. Res. Biogeosciences*, vol. 128, no. 1, p. e2021JG006471, Jan. 2023.

[5] J. Nieke and M. Rast, "Status: Copernicus Hyperspectral Imaging Mission for the Environment (CHIME)," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019, pp. 4609–4611.

[6] E. Ben-Dor *et al.*, "Using Imaging Spectroscopy to study soil properties," *Remote Sens. Environ.*, vol. 113, no. SUPPL. 1, pp. S38–S55, Sep. 2009.

[7] S. Chabrillat, E. Ben-Dor, J. Cierniewski, C. Gomez, T. Schmid, and B. van Wesemael, "Imaging Spectroscopy for Soil Mapping and Monitoring," *Surv. Geophys.*, vol. 40, no. 3, pp. 361–399, 2019.

[8] S. Priori, M. Zanini, V. Falcioni, and R. Casa, "Topsoil vertical gradient in different tillage systems : an analytical review in support of remote sensing predictive models," *Soil Tillage Res.*, vol. Submitted, 2023.

[9] A. Dupiau *et al.*, "MARMIT-2: An improved version of the MARMIT model to predict soil reflectance as a function of surface water content in the solar domain," *Remote Sens. Environ.*, vol. 272, p. 112951, Apr. 2022.

[10] A. Bablet *et al.*, "MARMIT: A multilayer radiative transfer model of soil reflectance to estimate surface soil moisture content in the solar domain (400–2500 nm)," *Remote Sens. Environ.*, vol. 217, pp. 1–17, Nov. 2018.

[11] B. Minasny *et al.*, "Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon," *Geoderma*, vol. 167–168, pp. 118–124, Nov. 2011.

[12] E. Vaudour *et al.*, "Temporal mosaicking approaches of Sentinel-2 images for extending topsoil organic carbon content mapping in croplands," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 96, 2021.

[13] R. Casa, S. Pignatti, S. Pascucci, F. Castaldi, and L. Marrone, "Estimation of agronomic soil properties from multitemporal PRISMA satellite imaging spectroscopy," in *European Conference on Precision Agriculture ECPA 2023*, 2023.

[14] S. Priori, L. Marrone, and R. Casa, "Hyperspectral PRISMA images and geophysical proximal sensing data fusion to map topsoil features : vineyard and arable land case studies," in *EGU General Assembly 2023*, 2023, pp. 5–6.