Topsoil properties estimation from PRISMA satellite images

Micol Rossini, Giulia Tagliabue, Cinzia Panigada, Chiara Ferrè, Luca Gallia, Luigi Vignali, Monica Pepe, Gabriele Candiani, Mirco Boschetti, Roberto Comolli, Roberto Colombo

2nd Workshop on International Cooperation in Spaceborne Imaging Spectroscopy
19 – 21 October 2022

Introduction

Scientific background and objectives

- The interest in soils has recently increased because they are facing pressures due to intensive agriculture, inappropriate land management, and climate change.
- Understanding how soil properties vary between and within agricultural fields allows for more efficient use of resources, improving agronomic and environmental management.
- The relationship between the soil reflectance spectrum in the optical domain and the topsoil properties have led to the development of promising data-driven methods of estimating soil properties.

OBJECTIVE

- Evaluation of the ability of **satellite PRISMA images** to estimate the soil organic carbon, clay, sand and silt content.

→ The recently launched (**PRISMA and EnMAP**) and upcoming (**CHIME**) hyperspectral satellites, featuring contiguous spectral data, are opening new opportunities for the accurate retrieval of topsoil properties using machine learning methods

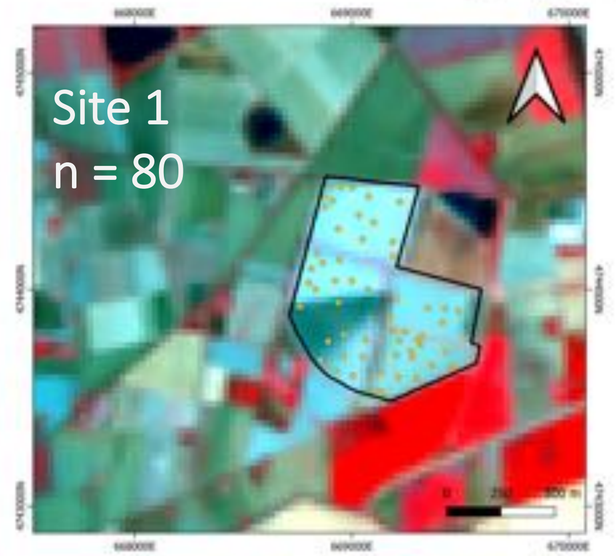
→ PRISMA (launched in March 2019) features 240 spectral wavebands from 400 to 2500 nm, providing the opportunity to accurately retrieve topsoil properties

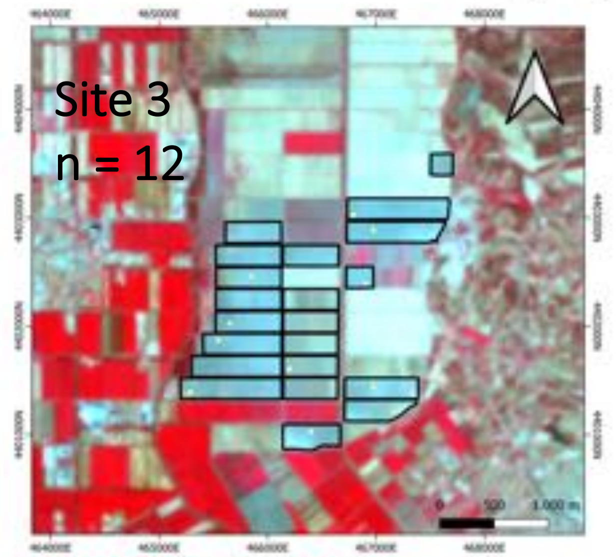
Credit: ASI

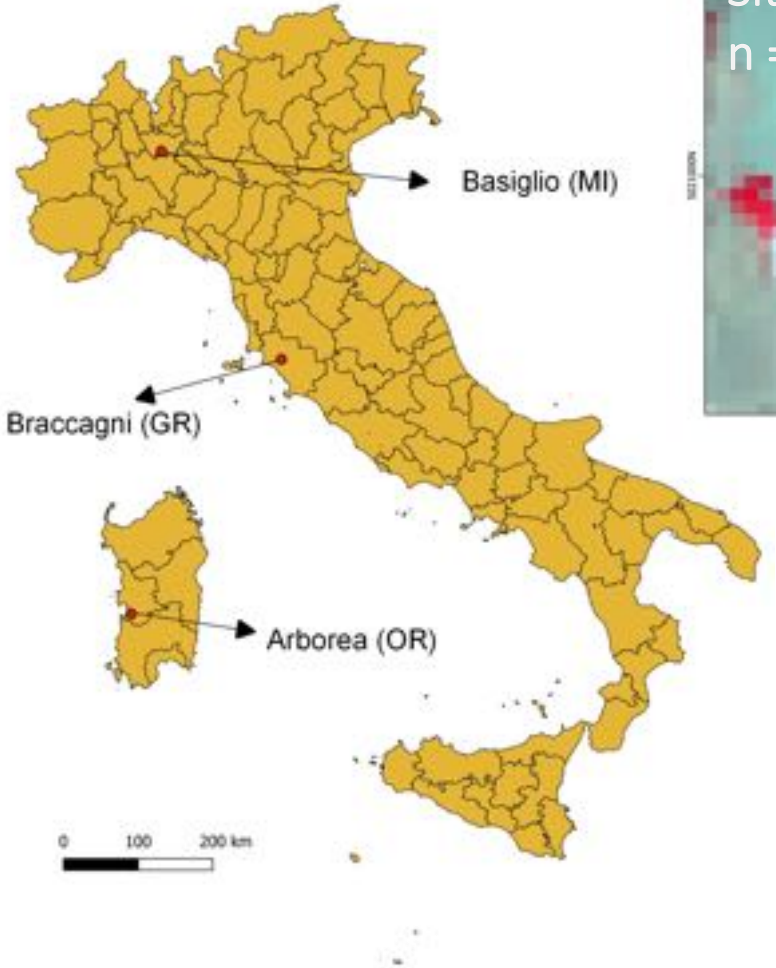
Study sites

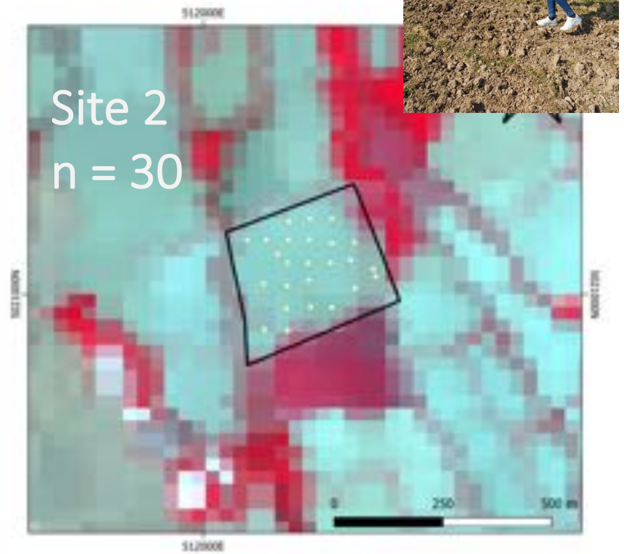
HyPlant June 2018

Spectral
Evolution
spectrometer
March 2022









Soil sampling
and laboratory
analysis

Soil sample data made available from BF in the framework of E-CROPS project

Data acquisition

Spaceborne imagery

→ **PRISMA** captured a spot image over

the **Basiglio area** on 24 April 2020

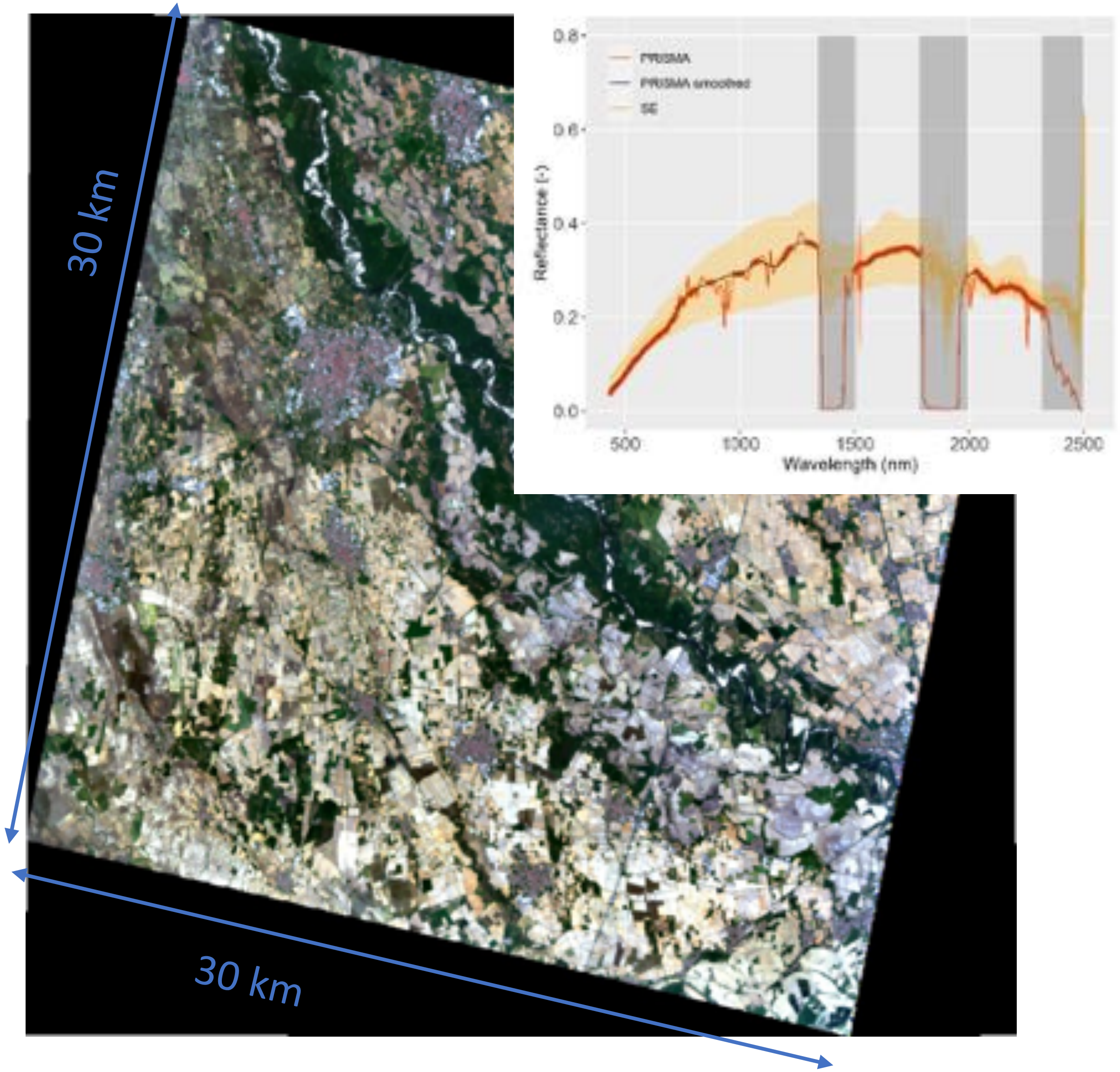
the **Braccagni area** on 21 June 2022

the **Arborea area** on 24 August 2019

→ The sensor is a push broom imaging spectrometer featuring **240 spectral bands** (400-2500 nm)

→ PRISMA has a swath width of 30 km and a **ground spatial resolution of 30 m**

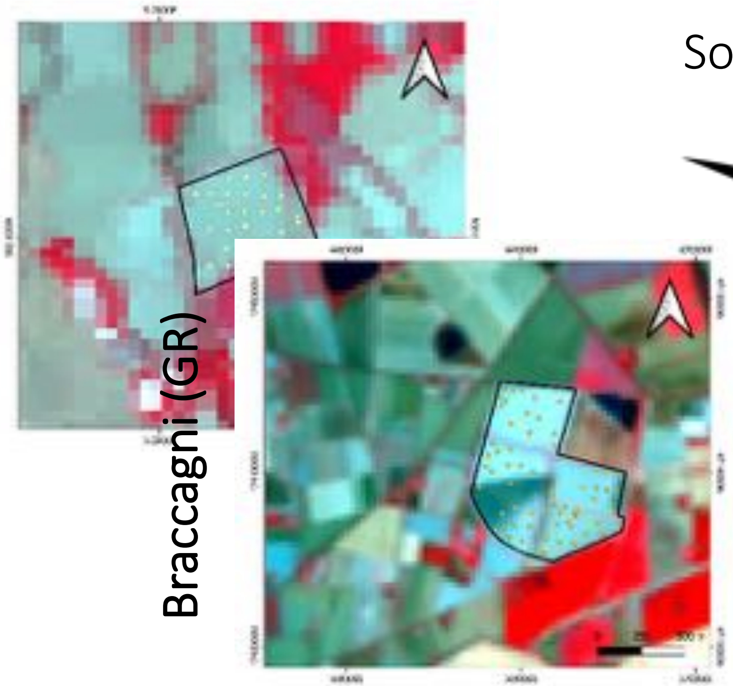
→ The L2D products (geocoded at-surface reflectance) were **pre-processed to obtain smooth spectra** (Tagliabue et al. 2022)



Methods
Workflow for the retrieval of topsoil parameters

Basiglio (MI)

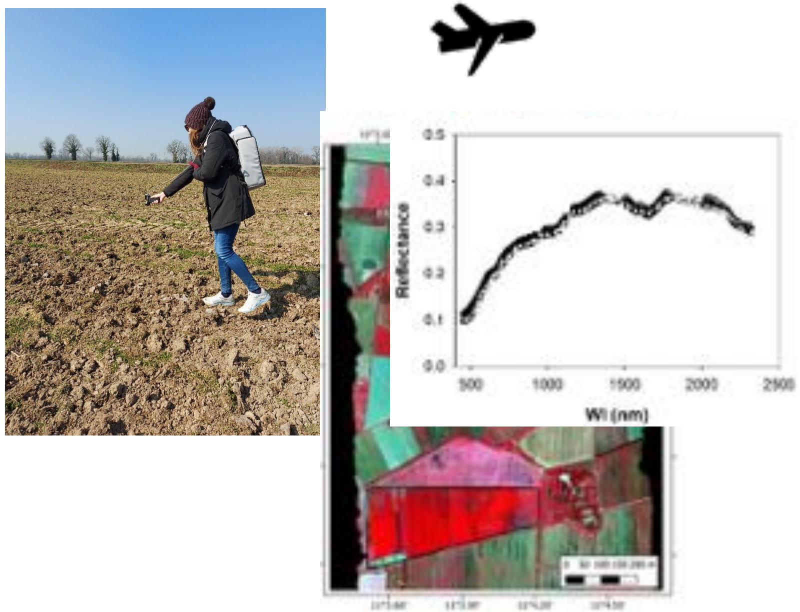
Braccagni (GR)



Soil sampling

Soil property values

Spectra at sampling points

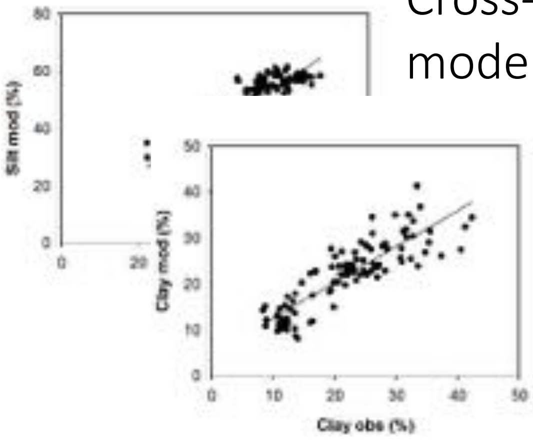


Machine learning (ML) training

ML algorithms
Least Squares Linear Regression
Partial Least Squares Regression
Random Forest
Neural Network
Support Vector Regression
Gaussian Process Regression

Dimensionality Reduction
PCA (5, 10, 15, 20)

Cross-val and model selection

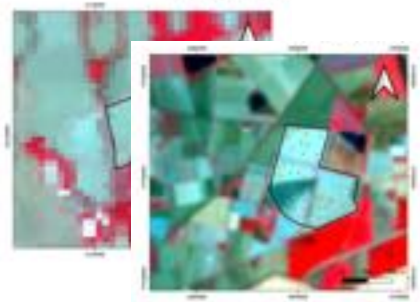


Methods

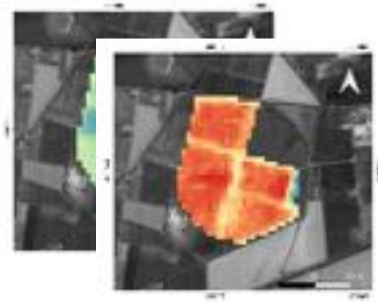
Workflow for the retrieval of topsoil parameters

Validation of the model | *sites 1 & 2*

Best performing ML algorithm



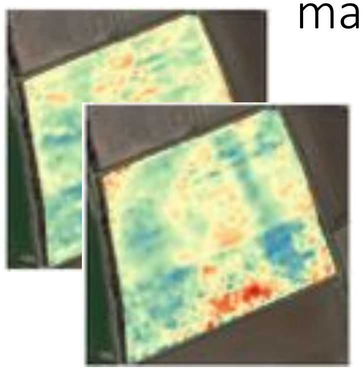
PRISMA images



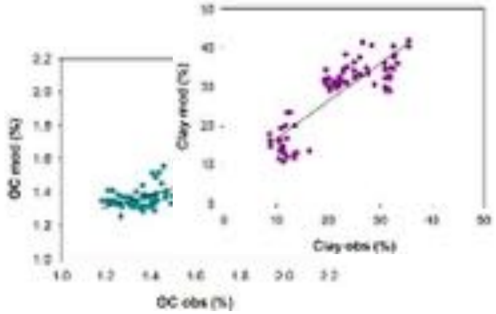
Soil property maps

Soil sampling

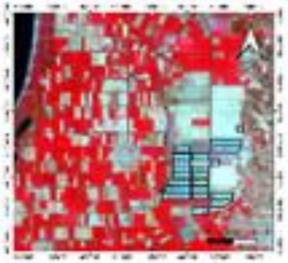
Geostatistic soil maps



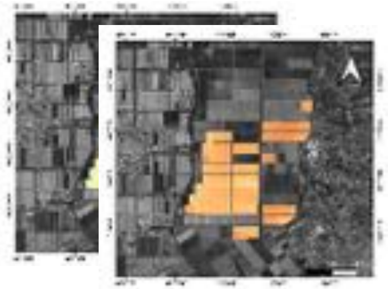
Validation



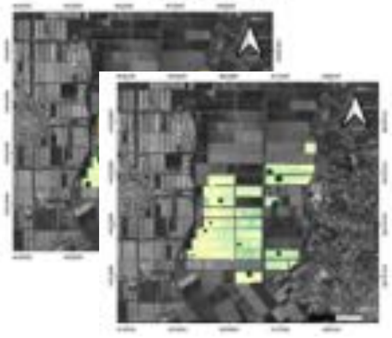
Exportability of the model | *site 3*



PRISMA image



Soil property maps



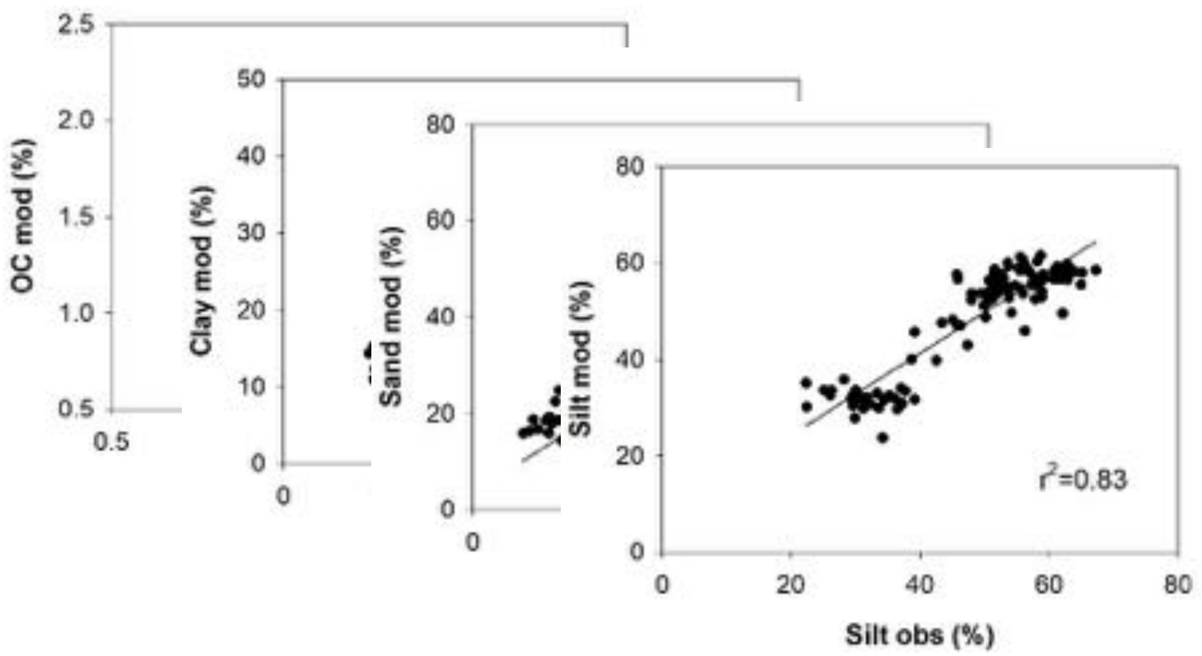
Evaluation of the maps with soil samples

Results

Model training with high spatial resolution data

- Models trained using as input the first 10 synthetic bands obtained through PCA.
- The prediction accuracy of the airborne MLRA models built using the high spatial resolution dataset (n=106) was evaluated using a leave-one-out validation.
- The Least Square Linear Regression algorithm provided the best models for the retrieval of the four soil properties.

Parameter	MLA	r^2	RMSE	nRMSE
SOC %	LSLR	0.4	0.17	14.3
Clay %	LSLR	0.75	4.15	12.2
Sand %	LSLR	0.88	6.07	9.9
Silt %	LSLR	0.83	5.02	11.2

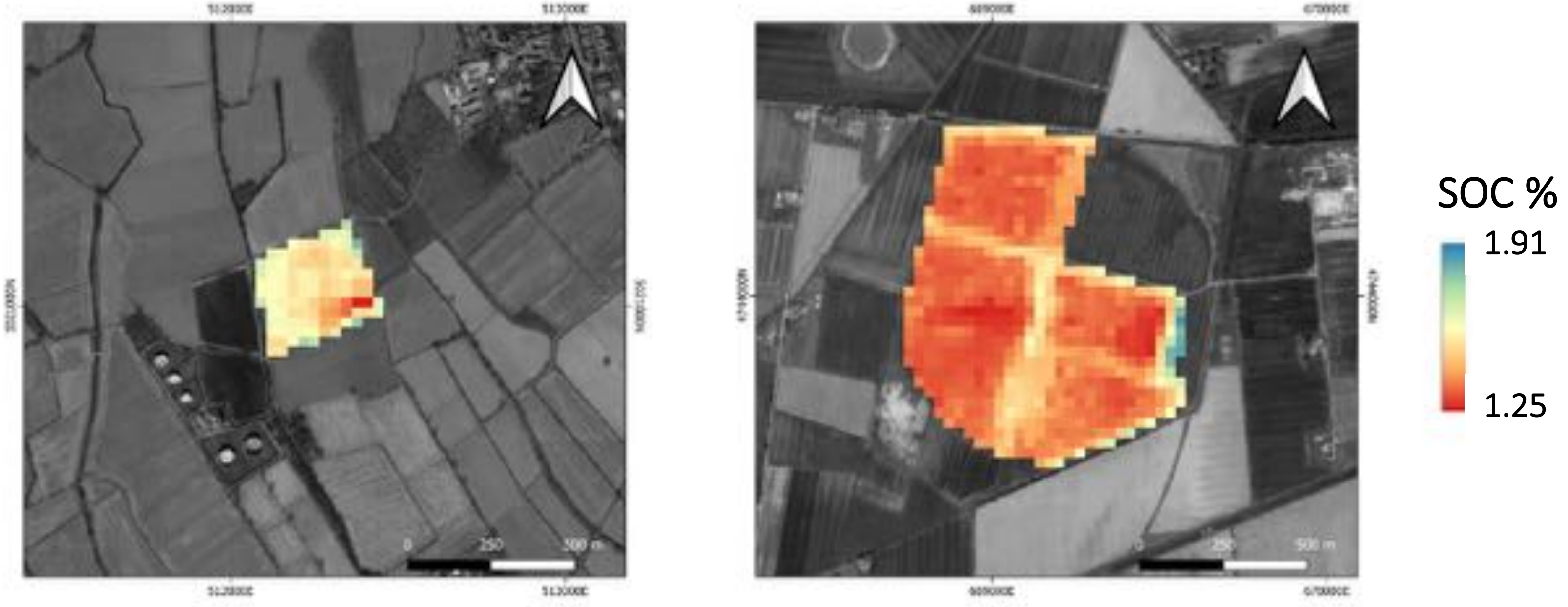


Results

Validation of PRISMA topsoil parameter maps

- Best performing model applied to hyperspectral PRISMA imagery at the Basiglio and Braccagni study sites.
- Comparison between the topsoil parameters estimated using PRISMA data and the values obtained from the geostatistic maps (n = 73).

Parameter	r^2	RMSE	RPD	RPIQ
SOC %	0.45	0.13	1.37	1.86
Clay %				
Sand %				
Silt %				

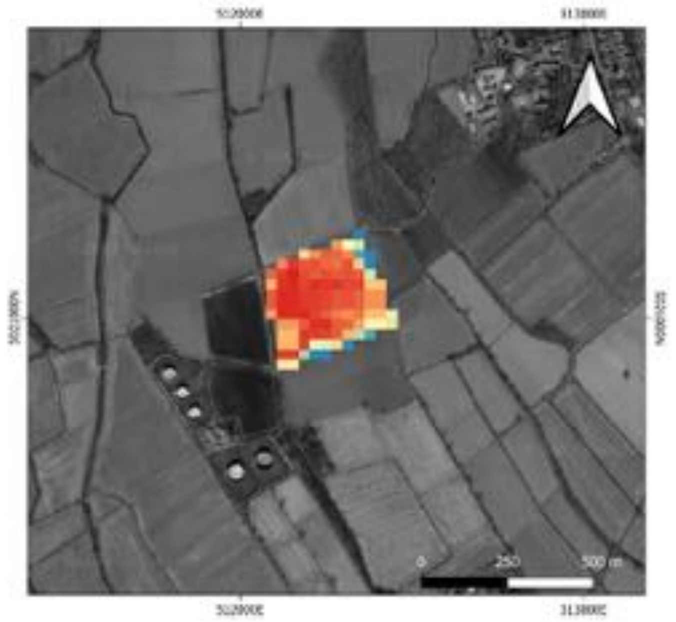


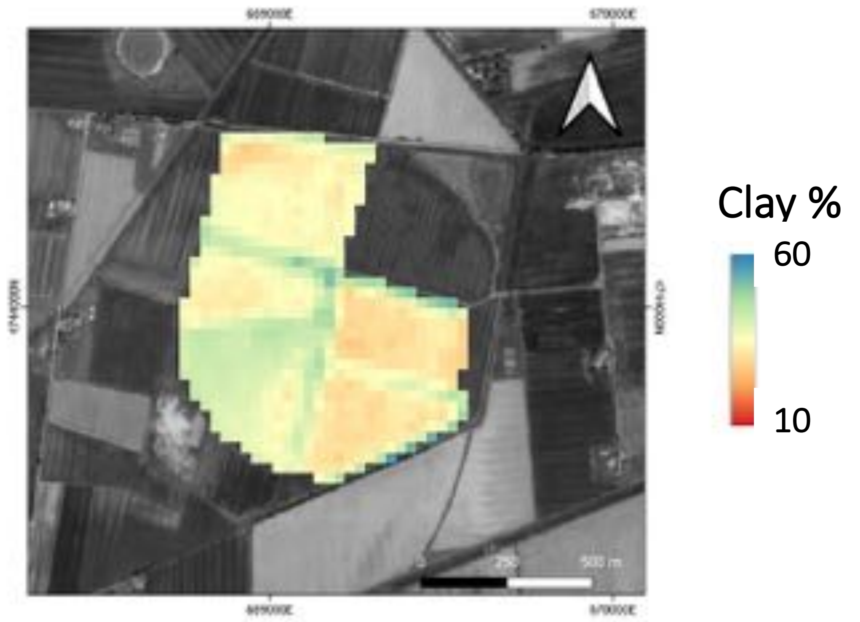
Results

Validation of PRISMA topsoil parameter maps

- Best performing model applied to hyperspectral PRISMA imagery at the Basiglio and Braccagni study sites.
- Comparison between the topsoil parameters estimated using PRISMA data and the values obtained from the geostatistic maps (n = 73).

Parameter	r^2	RMSE	RPD	RPIQ
SOC %	0.45	0.13	1.37	1.86
Clay %	0.74	7.68	1.05	2.26
Sand %				
Silt %				



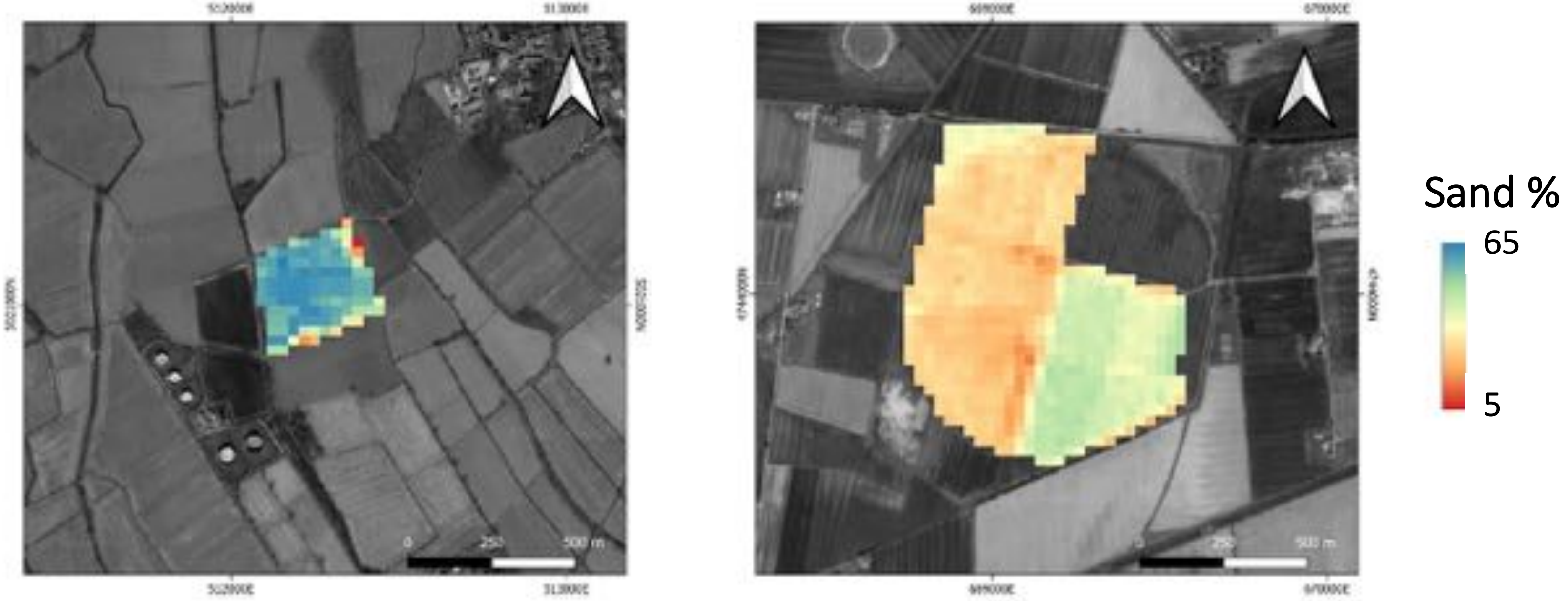


Results

Validation of PRISMA topsoil parameter maps

- Best performing model applied to hyperspectral PRISMA imagery at the Basiglio and Braccagni study sites.
- Comparison between the topsoil parameters estimated using PRISMA data and the values obtained from the geostatistic maps (n = 73).

Parameter	r^2	RMSE	RPD	RPIQ
SOC %	0.45	0.13	1.37	1.86
Clay %	0.74	7.68	1.05	2.26
Sand %	**0.86**	**16.78**	**1.13**	**2.24**
Silt %				

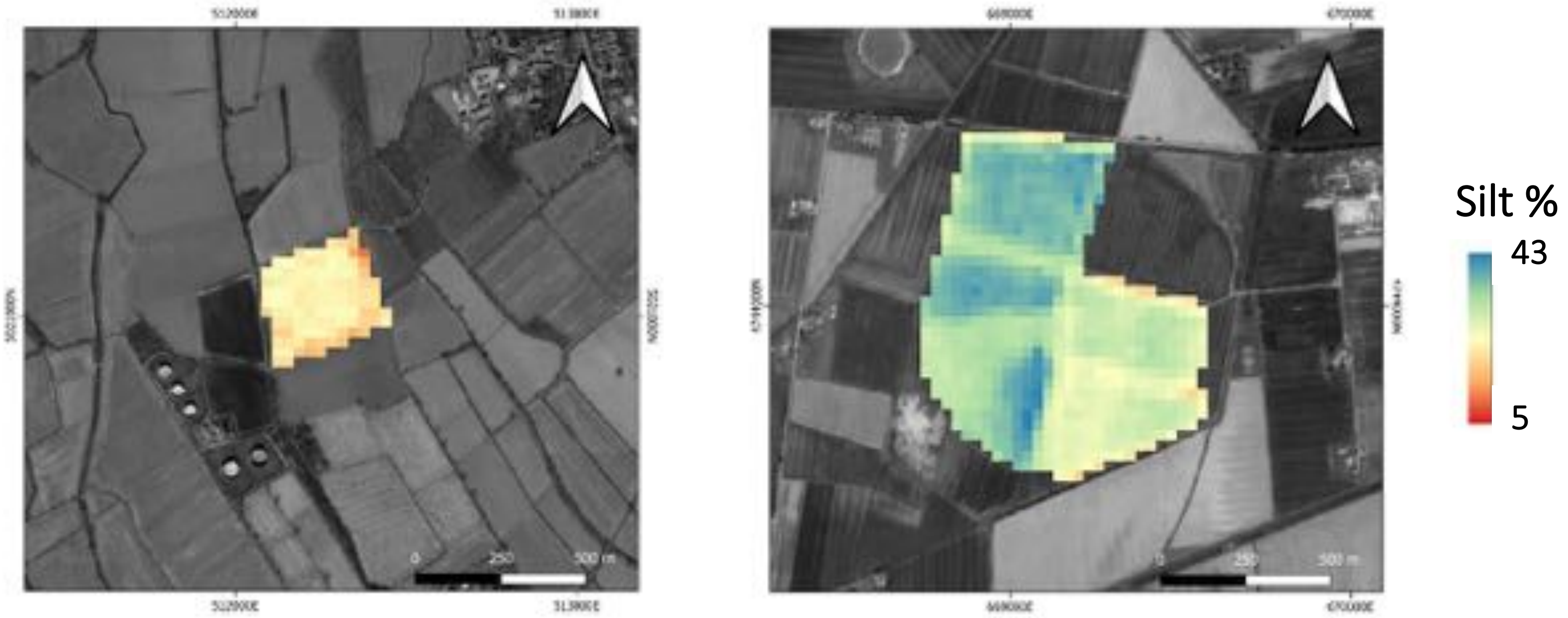


Results

Validation of PRISMA topsoil parameter maps

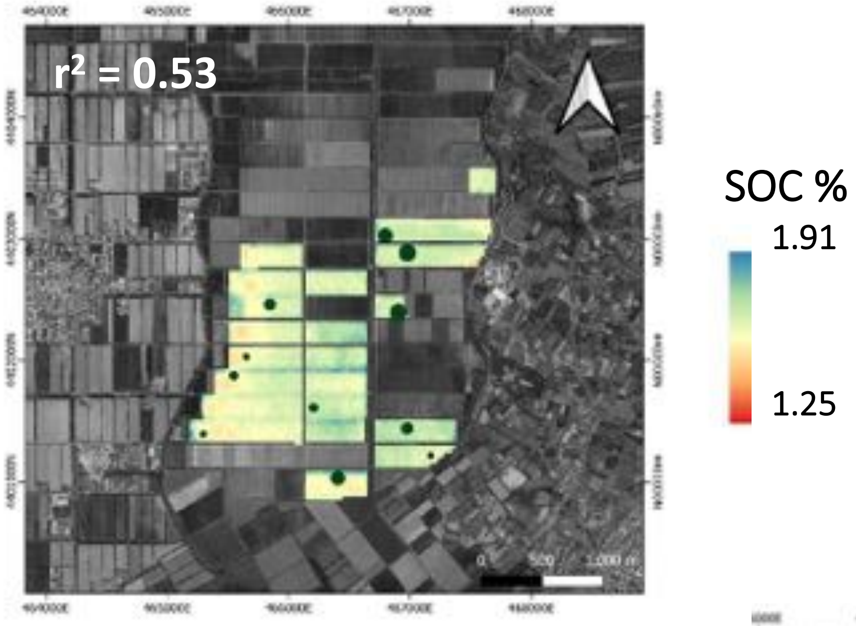
- Best performing model applied to hyperspectral PRISMA imagery at the Basiglio and Braccagni study sites.
- Comparison between the topsoil parameters estimated using PRISMA data and the values obtained from the geostatistic maps (n = 73).

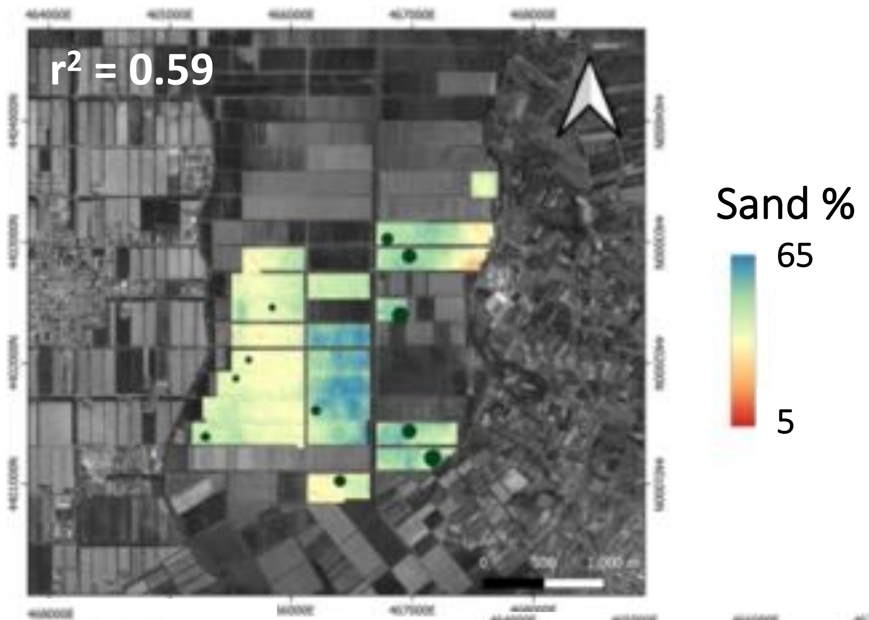
Parameter	r^2	RMSE	RPD	RPIQ
SOC %	0.45	0.13	1.37	1.86
Clay %	0.74	7.68	1.05	2.26
Sand %	0.86	16.78	1.13	2.24
Silt %	0.46	22.7	0.54	0.94



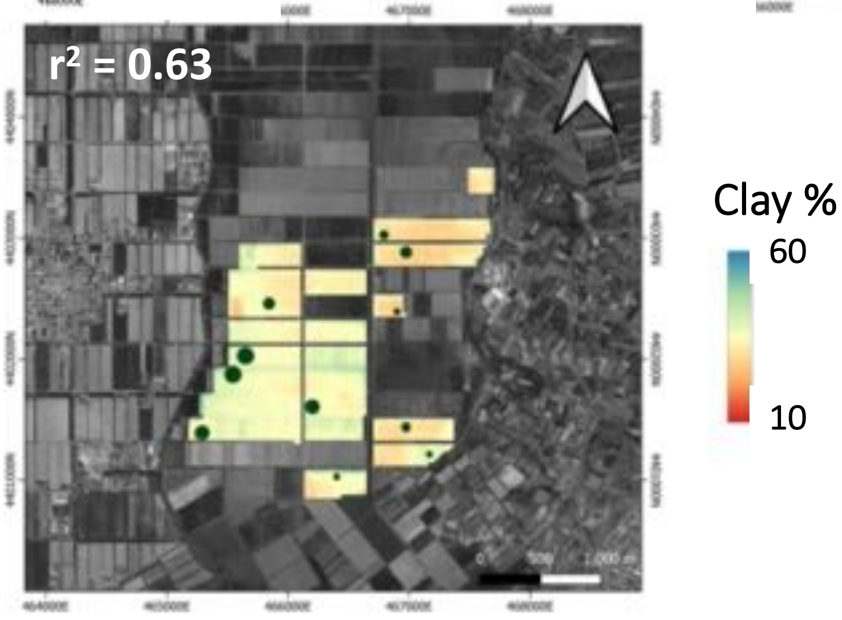
Results

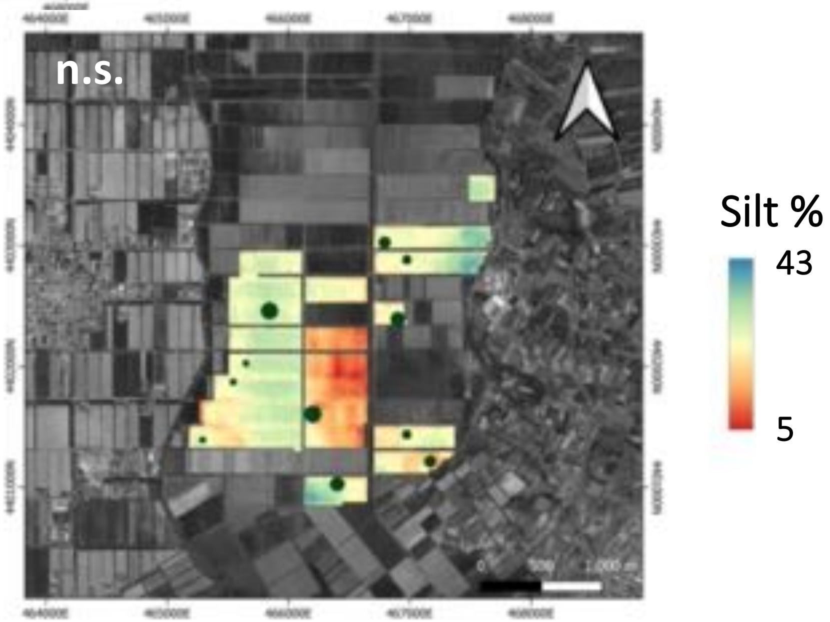
Evaluation of model exportability





- Best performing model applied to hyperspectral PRISMA imagery at the Arborea study site.
- Comparison between the PRISMA topsoil maps and soil samples.





Conclusions
Take home messages

- We evaluated the ability of satellite PRISMA images to estimate the soil organic carbon, clay, sand and silt content over bare soils in croplands.
- Machine learning algorithms were trained using high spatial resolution hyperspectral data. The best performances were obtained using LSLR with cross-validated r^2 of 0.4 for OC, 0.75 for clay, 0.88 for sand and 0.83 for silt.
- SOC estimation results were not very good, probably because of the limited range of variation of this variable within the fields.
- The developed models were successfully validated using PRISMA images collected over two experimental sites in 2020 and 2022 demonstrating their robustness on an independent image dataset.
- The performances of the models decrease when applied on different geographical areas (different management of the soil). More investigations are needed to evaluate model exportability.
- The results obtained demonstrate that the retrieval of soil properties from space using ML algorithms is feasible, paving the way for future operational algorithms for topsoil mapping from hyperspectral satellites (PRISMA, CHIME, EnMap).

Thank you!

micol.rossini@unimib.it