UNIVERSITÁ DEGLI STUDI DEL MOLISE

Department of Agricultural, Environmental and Food Sciences

---

PhD Course in

AGRICULTURE TECHNOLOGY AND BIOTECHNOLOGY

(CURRICULUM: SUSTAINABLE PLANT PRODUCTION AND PROTECTION)

CYCLE XXXIII

Related Disciplinary Scientific Sector: AGR/14 (Pedology)

PhD Thesis

# Integration Vis-NIR Spectroscopy and satellite Sentinel-2 data to assess soil organic carbon pool of an organic farm in South of Italy

Coordinator of the PhD Course: Prof. Giuseppe Maiorano

Supervisor: Prof. Claudio Colombo

coSupervisor: Dr. Erika Di Iorio

PhD Student: Luana Circelli

160424

---

Academic Year 2019/2020

## Abstract

The soil plays an important role in the global carbon cycle and the balance between the rates of organic carbon input and output determines at large scale the amount and the turnover of soil organic carbon (SOC). Thus, soil can act as both carbon store and supplier of carbon emissions increasing atmospheric greenhouse gases (GHGs), having a great potential in climate change mitigation. Changes in land use are those particularly involved in SOC accumulation and in loss of carbon dioxide ($CO_2$) in the atmosphere as GHG. With this purpose, sustainable land management are necessary to be applied for carbon storage and sequestration. However, assessing the composition of SOC is important to understand the mechanisms through which soil C pools can be directly linked with its potential to sequester atmospheric $CO_2$. The study of the SOC composition span from the micro to the macro-scale, and the application of Vis-NIR spectroscopy technique can couple some positive aspects of the resolution of the fine scale with a wide and easy applicability. In this framework, as a rapid technique for measuring SOC based on the diffuse soil reflectance, the application of Vis-NIR spectroscopy and integration with satellite Sentinel-2 data were applied.

With the aim to analyze the organic and inorganic C pools, the soils of an organic farm in South of Italy were chosen. In this farm, the principles of conservation agriculture are adopted in apple cultivation. The organic and inorganic C pools were measured on 56 georeferenced soil samples, for a total study area of about 3 hectares, applying different methods, from dry combustion at high temperature, CNS analyzer and furnace loss on ignition for total and organic C, to wet oxidation procedures, also for soluble forms of organic C. This differentiation in methods allowed making a distinction between the more stable (recalcitrant) organic C forms in the soil from the more labile ones and more subject to mineralization phenomena. Spatial regression models allowed taking into account the role of topographical parameters and other soil properties on dynamics of organic and inorganic C pools. Vis-NIR spectroscopy, as a proximal sensing technique, was used to model calibrations and validations for the prediction of the investigated soil C fractions. From the combinations between different spectral preprocessing and regression models, second derivative transformations and partial least squares (PLS) regression have achieved excellent results, made by the average values of $R^2$=0.85 and RPD=2.5 for soil organic C predictions, and $R^2$=0.75 and RPD=1.7 for soil inorganic C. The evaluation gap of soil C pools from field to farm scale was improved by the

satellite Sentinel-2 data in Vis-NIR predictive models. The obtained accurate results proved that remote sensing could be a powerful tool to recognizing the various soil C pools, especially those most exposed to the risk of degradation and erosion. In addition, the combination between soil laboratory analysis and Vis-NIR spectroscopy and satellite data can be very useful with a precision agriculture approach.

## Acknowledgements

As an important contribution to the drafting of this dissertation, first I want to thank Prof. Claudio Colombo, Supervisor of my PhD, who has allowed me over these years to deepen and appreciate more and more Soil Science.

I also want to thank Dr. Erika Di Iorio, for always advising and guiding me during this experience, and all the people who in various ways have contributed to making me grow professionally.

I thank Prof. Giuseppe Maiorano, Coordinator of PhD Programme, for allowing all the PhD activities to be carried out in the best possible way.

Lastly, I would like to thank the Melise farm and all its collaborators for being constantly available during the field sampling and sharing of activities.

# Contents

**List of Tables** **VII**

**List of Figures** **VIII**

**Introduction** **1**

1.1 The potential carbon sequestration of the soil ..... 1

1.1.1 Knowledge about the organic carbon pools of the soil ..... 3

1.2 Large scale and fast techniques for the assessment of soil organic carbon: proximal and remote sensing ..... 9

1.2.1 Sentinel-2 as a powerful tool for validating soil organic carbon pools ..... 13

1.3 Aim of the thesis ..... 15

**Materials and Methods** **16**

2.1 Study area ..... 16

2.1.1 Agronomic, pedological and climatic characterization of the area ..... 17

2.1.2 Experimental design ..... 20

2.2 Soil physico-chemical analysis ..... 20

2.2.1 Carbon fractionation ..... 21

2.3 Soil biological analysis ..... 22

2.4 Statistical analysis of soil properties ..... 23

2.5 Soil spectra analysis ..... 24

2.5.1 Data preparation and preprocessing transformations ..... 25

2.5.2 Prediction models development ..... 26

2.6 Sentinel-2 data collection ..... 29

2.7 Geostatistical analysis ..... 31

**Results and Discussion** **32**

3.1 Descriptive statistics of soil properties ..... 32

3.1.1 Carbon fractions ..... 32

3.1.2 Soil properties affecting the organic and inorganic carbon fractions ..... 34

3.1.3 Soil properties affecting the labile organic carbon fraction and biological properties ..... 39
3.1.4 The organic carbon accumulation and soil properties ..... 39
3.1.5 Results of principal component analysis ..... 41
3.2 Spatial prediction of soil carbon pools ..... 43
3.3 Predictions of soil carbon pools using Vis-NIR data ..... 52
3.3.1 Vis-NIR spectra characterization ..... 52
3.3.2 Soil carbon pools affecting Vis-NIR spectra ..... 55
3.3.3 Performances of carbon pools predictions ..... 59
3.4 Validation of Vis-NIR predictions using external remote sensing Sentinel-2 data ..... 67
3.4.1 Spatial prediction of Sentinel-2 carbon pools from field to farm-scale ..... 70

**Conclusions** **75**

**A Statistics details of measured soil carbon pools** **78**

**List of References** **82**

## List of Tables


| | | |
|---|---|---|
| 1.1 | Comparison of methods used for determination of total C in soils ……... | 8 |
| 2.1 | Specifications of a Sentinel-2 Multispectral Instrument sensor ……... | 29 |
| 2.2 | Main characteristics of the studied scenes ……... | 30 |
| 3.1 | Basic statistics of the analyzed soil variables ……... | 35 |
| 3.2 | Pearson correlation matrix between soil properties ……. | 37 |
| 3.3 | Performances of application of geostatistical models …….. | 43 |
| 3.4 | Performances of calibration of soil carbon pools predictive models …… | 61 |
| 3.5 | Performances of validation of soil carbon pools predictive models ……. | 62 |
| 3.6 | Values of explained X total variance from Sentinel-2 spectra …….. | 68 |

## List of Figures


| | | |
|---|---|---|
| 1.1 | Plant C input and loss pathways in soil ……..…..... | 1 |
| 1.2 | Constituents of soil carbon pool ……... | 4 |
| 1.3 | Mechanisms of stabilization of soil organic matter ……..…… | 6 |
| 1.4 | Electromagnetic spectrum in Vis-NIR-SWIR ……. | 9 |
| 1.5 | Schematic diagrams of proximal and remote sensing …….. | 10 |
| 1.6 | Spectra analysis for reflectance spectroscopy …….. | 13 |
| 2.1 | Geographical location and altimetric map of Melise farm ……... | 16 |
| 2.2 | Characterization of soil profile sampled in studied area ……... | 18 |
| 2.3 | Climatic characterization of Fornelli area …… | 19 |
| 2.4 | Cadastral map of Melise farm and sampling design ……. | 19 |
| 2.5 | Schematic description of the soil carbon fractionation procedure …… | 22 |
| 2.6 | Geometry of the integration sphere in the spectrophotometer …….. | 24 |
| 2.7 | Spectra preprocessing of all soil samples ……. | 27 |
| 3.1 | Box plots for the partitioning of soil carbon pools ……... | 33 |
| 3.2 | Correlations between measured soil organic carbon fractions ……. | 34 |
| 3.3 | Soil texture contents of samples ……... | 40 |
| 3.4 | Plots resulting from principal component analysis …….. | 42 |
| 3.5 | Experimental and fitted model semivariograms of soil carbon pools ……... | 44 |
| 3.6 | Experimental and fitted model semivariograms of soil properties ……... | 45 |
| 3.7 | Spatial predictive maps of soil carbon pools …… | 48 |
| 3.8 | Spatial predictive maps of soil properties …… | 49 |
| 3.9 | Slope and aspect, and calculated NDVI and NDWI for studied area …… | 51 |
| 3.10 | The mean soil spectrum and standard deviation values …… | 53 |
| 3.11 | Pearson's correlation coefficients between spectral reflectance …….. | 56 |
| 3.12 | Score plot resulting from PCA ……. | 57 |
| 3.13 | Correlation loadings line plot resulting from PCA ……... | 58 |
| 3.14 | Coefficient of determination in validation for soil organic C fractions …… | 63 |
| 3.15 | Coefficient of determination in validation for soil inorganic C fractions ……. | 64 |
| 3.16 | Scatter plots of Vis-NIR predicted values of soil carbon pools …… | 65 |

3.17 Summarizing of VIP profiles from PLSR …… 66
3.18 Raw reflectance laboratory and Sentinel-2 spectra …….. 67
3.19 Correlations and frequency distribution of Sentinel-2 organic C fractions …….. 69
3.20 Correlations and frequency distribution of Sentinel-2 inorganic C fractions ….. 70
3.21 Spatial predictive soil maps of Sentinel-2 organic carbon fractions ……. 71
3.22 Spatial predictive soil maps of Sentinel-2 inorganic carbon fractions …….. 72
3.23 Experimental and fitted model semivariograms of Sentinel-2 values …….. 73

A.1 Histogram plot of $TC_{CNS}$ soil parameter ……..…… 78
A.2 Histogram plot of $OC_{CNS}$ soil parameter …….. 79
A.3 Histogram plot of $IC_{CNS}$ soil parameter …… 79
A.4 Histogram plot of $TOC_{WB}$ soil parameter …… 80
A.5 Histogram plot of HC soil parameter ……... 80
A.6 Histogram plot of LOI soil parameter …….. 81
A.7 Histogram plot of $CaCO_3$ soil parameter ……. 81

# Chapter 1
# Introduction

## 1.1 The potential carbon sequestration of the soil

The soil plays an important role in the global carbon cycle and the balance between the rates of organic carbon input and output determines at large scale the amount and the turnover of soil organic carbon (SOC) (Lal, 2004, 2008; Krull et al., 2004; Doetterl et al., 2015). The carbon input is the sum of a large variety of organic carbon compounds added to the soil, crop residues, roots that die off, live fauna, organic carbon in manure and wastes. As recently suggested by Navarro-Pedreño et al. (2021), organic carbon gain in soil involves storage and a possible reduction of $CO_2$ in the atmosphere. When organic compounds enter the soil, they can be decomposed partially or totally by microbial life to extract chemical energy from them, which, under aerobic conditions, results in the emission of $CO_2$. On the contrary, carbon sequestration leads to storage of the organic compounds in the soil, through the removal of $CO_2$ from the atmosphere (Figure 1.1).

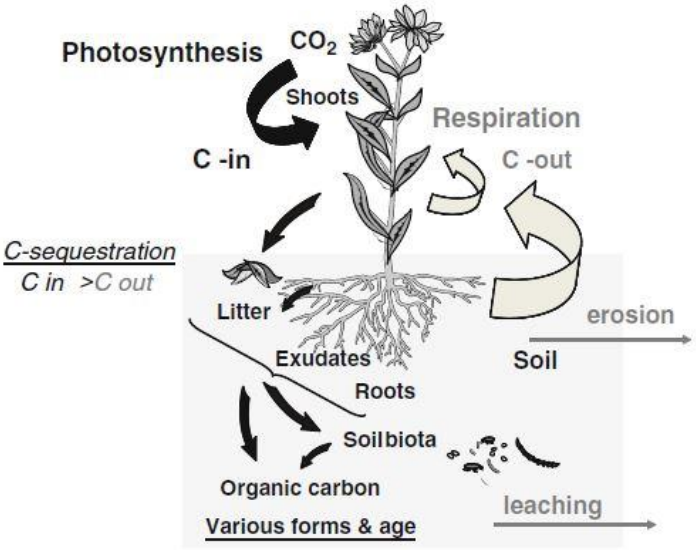


Figure 1.1 Plant C input and loss pathways in soil: main input photosynthesis, main output respiration, erosion and leaching of C, as well as harvest, herbivory, emissions of volatile organic compounds (De Deyn, 2015).

The carbon sequestration, described as long-term soil response, was considered by Bernoux et al. (2006) as "the result for a given period of time and portion of space of the net balance of all GHG expressed in C-$CO_2$ equivalent or $CO_2$ equivalent computing all emissions sources at the soil–plant–atmosphere interface, but also all the indirect fluxes". Thus, soils can act as both carbon store and supplier of carbon emissions increasing atmospheric greenhouse gases (GHGs), having a great potential in climate change mitigation. Moreover, the role of soils is essential in global negotiations ($CO_2$ market) as carbon storage mechanisms to mitigate climate change, being an important part of the carbon footprint over the life stages of a product (Navarro-Pedreño et al., 2021).

SOC and its potential to become a "managed" sink for atmospheric carbon dioxide ($CO_2$) has been widely discussed in the scientific literature (Kirschbaum, 2000; Post and Kwon, 2000; Guo and Gifford, 2002; Post et al., 2004; Smith, 2008; Chabbi and Rumpel, 2009; Luo et al., 2010). The factors that most affect the potential of carbon storage of the soil consist of the environmental conditions (climatic factors), soil management and the type of soil and transformation processes, including biological interactions (von Lützow and Kogel-Knabner, 2009). These factors include soil temperature, soil moisture and water saturation, texture, topography, salinity, acidity, vegetation, and biomass production (Bot and Benites, 2005; Magdoff and Van, 2010). Above all, land management practices are those that have the greatest influence on carbon storage potential (Stockmann et al., 2013; Paustian et al., 2016). SOC is a component part of soil organic matter (SOM), and in order to assess the SOC relate to SOM, it is generally accepted that the factor 1.72 estimates SOC 58% of the SOM total. SOM is different from SOC, because SOM includes together with organic carbon all the elements H, N, O, etc. which are also part of organic compounds (Swift, 1996). In general, SOC is extremely dynamic, because its highly reactive, its used as a source of energy for all microorganisms and other biota in the soil, and is preferentially removed by erosional processes because it has low density and is located in vicinity of the soil surface (Lal, 2015). Because of this dynamic, sustainable land management practices are desirable to be applied for carbon store and sequestration.

Changes in land uses are those particularly involved in soil organic carbon contents. As suggested by Jafarian and Kavian (2013), the soils in the native forest ecosystems have a significantly greater organic C content than those in the dry lands and grazing lands, due to increased (above and belowground) biomass and reduced litter decomposition rates. Guo and Gillford (2002) highlighted the influence of land use changes on soil C stocks, demonstrating that changing land use from cropland to pasture or cropland to permanent forest result in the greatest gains of SOC. On the contrary, Minasny et al. (2011) after long-term investigation on

dynamics of SOC found that changes in land use from forest to plantation and cultivated land resulted in increasing of total SOC, denoting thus that human interventions to increase plant production through fertilizer application and other management can also positively affect SOC contents.

Different authors (Bian et al., 2013; Stockmann et al., 2013; Kämpf et al., 2016; García-Díaz et al., 2018) agree that no-till, conservation tillage, groundcover and cover crops integrated in grain, forage and agro-forestry systems, and grassland management are the best SOM management practices, to enhance carbon sequestration and ensuring long term productivity.

Nevertheless, as suggested by Yang et al. (2019), the total soil organic carbon cannot reflect changes of the introduction of sustainable land practices because it often takes years before shifts in agricultural management show detectable variations in SOC content. It is essential then understand how SOC is lost or stabilized in soil into different functional pools depending on their varying residence time (Sahoo et al., 2019). As considered by Haynes (2005), labile pool (active pool) is the most sensitive pool available relatively in small proportion as it is easily affected by fluctuation in environmental conditions. On the contrary, the non-labile pool (passive pool) is more stable and recalcitrant fraction of SOC forming organic-mineral complexes with soil mineral and gets decomposed slowly by microbial activity (Wiesenberg et al., 2010). Indeed, in the more recent years, researches and knowledge about different soil organic carbon pools have received more attention due to their sensitivity to management practices than bulk SOM or SOC stock.

### 1.1.1 Knowledge about the organic carbon pools of the soil

Carbon (C) storage is an important ecosystem function of soils that has gained increasing attention in recent years due to its interactions with the earth's climate system (FAO and ITPS, 2015). Now, it is widely accepted that the soil is a major C reservoir that holds more carbon than is contained in the atmosphere and terrestrial vegetation combined, and all three of these reservoirs are in constant exchange. Understanding the composition of SOC is important because soil OC comprises many different organic compounds and structures, each interacting uniquely with the soil component and decomposing rate. While the physical forms can contain similar classes of organic compounds, such as phenolics, proteins, lipids or carbohydrates, they may differ in origin from plant residues or in situ microbial remains (Tate, 1987). Conversely, they can be of similar origin but be present in different parts of the organic matter and represent different stages of decomposition, such as lignin in the more labile part, poly-phenolic compounds in the humic part or char-carbon phenolic residues in the more resistant part

(Stevenson, 1994). In many soils SOC, which contains roughly 55-60 percent C by mass, comprises most or all of the C stock, especially when there is a scarce carbonate concentration. Generally, SOC is composed of plant litter compounds as well as of microbial decomposition products. SOC is thus a complex biogeochemical mixture derived from organic material in all stages of decomposition (von Lützow et al., 2006; Paul, 2014). Due to microbial degradation and mineralization to $CO_2$, the majority of plant litter compounds added to soil remain for a relatively short time (from a few days to a few years). However, some organic matter compounds may persist in the soil for decades or centuries or even for millennia (Paul et al., 1997; von Lützow et al., 2006). Considered as one of the key indicators of soil fertility and quality, SOC is comprised of diverse and heterogeneous organic substances that can then significantly influence the physical and biochemical reactions that occur in soils (Cambule et al., 2012; Dhillon et al., 2017). SOC content, considered as a good indicator of a healthy soil system, is related to a wide range of soil properties and processes, such as soil structure, water-retention capacity, nutrient cycling, and biological activity (de la Paz Jimenez et al., 2002).

Over the last two decades, research on SOC has focused largely on their chemical structure and the effects of inorganic colloids on the polymerization of single organic molecules (Schnitzer and Schulten, 1992; Liu and Huang, 2002). However, in the last years conceptual SOC models changes and experimental advances have demonstrated that molecular structure alone does not control SOC stability (Schmidt et al., 2011). SOC indeed, can be protected in the soil matrix through physical stabilization processes, as well as by inherent "chemical

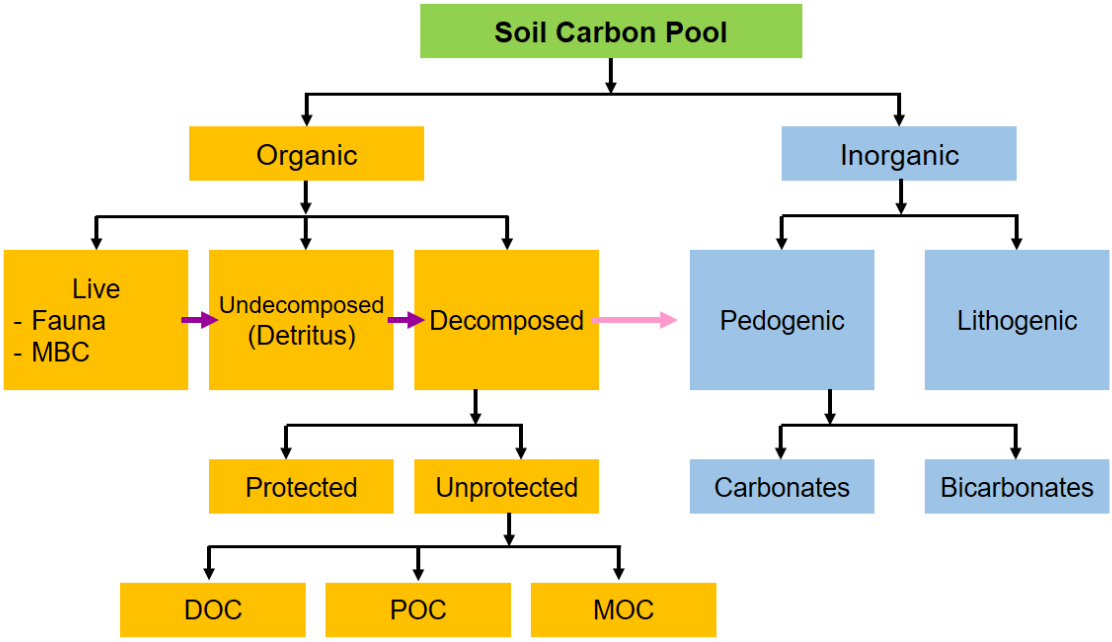


Figure 1.2 Constituents of soil carbon pool (DOC, dissolved organic C; POC, particulate organic C; MOC, mineral organic C) (adapted from Lal, 2017).

recalcitrance". According to Six et al. (2004), the main physico-chemical stabilization processes are the protection within aggregates which translates to spatial inaccessibility of soil microbes to organic compounds and a limitation on oxygen availability and, interactions with mineral surfaces and metal ions. Now, SOC is conventionally divided into at least three C pools, the "active pool" (root exudates, rapidly decomposed components of fresh plant litter) with turnover rates of years, the intermediate or "slow pool" with turnover rates of decades and the "passive pool" (stabilized organic matter due to chemical or physical mechanisms) with turnover rates of centuries to millennia (Trumbore, 2009).

It is widely accepted that soil C to contribute to climate regulation by sequestration, needs to be in a stable, non-labile form that will not be susceptible to losses should the system be perturbed by a change in tillage or by climatic changes that increase microbial activity (Lal, 2004; Powlson et al., 2012). Indeed recently, dissolved organic matter (DOM) has been defined as the most dynamic and bioavailable fraction of SOC, responsible for stimulating soil microbial activity to promote the decomposition of SOM (Liu et al., 2019). Different authors have deepened their researches on labile forms of organic carbon, to comprehend the mechanisms through which soil C pools can be directly linked with its potential to sequester atmospheric $CO_2$. Solubility in water is one attribute of readily bioavailable SOC. Different extractants are used to obtain DOM, ranging from cold water to aqueous solutions of different ionic strength to simulate the soil solution, since there are no standard methods for measuring it. Wen et al. (2016), Tokarski et al. (2020) and Zhang et al. (2020) isolated two different organic carbon fractions, both as labile C pools or active pool according to Zimmermann et al. (2007), which are water-soluble organic carbon (WSOC) and particulate organic carbon (POC), obtained with extraction with soil to water solution and density solution, respectively. They found that combining crop straw residues with chemical fertilization could be an important strategy for increasing labile fractions of soil organic matter, which in turn significantly enhance soil macroaggregate content and then OC storage. Sarkhot et al. (2011) obtained the same results extracting available C forms with soil to water solution heated at 80°C to facilitate the dissolution of labile organic carbon fractions in water.

On the other hand, more stable and recalcitrant organic carbon fractions have suggested to be important for long-term sequestration of SOC. Different authors (Helfrich et al., 2007; von Lützow et al., 2007; Álvaro-Fuentes et al., 2008; Poeplau et al., 2013 and 2018) managed to isolate two slow cycling organic material, SOC attached to sand and clay grains in stable aggregates without being chemically resistant, and a chemically resistant fraction. The first one is obtained treating soil in a density solution, as the heavy fraction (HF) of POM, and the second

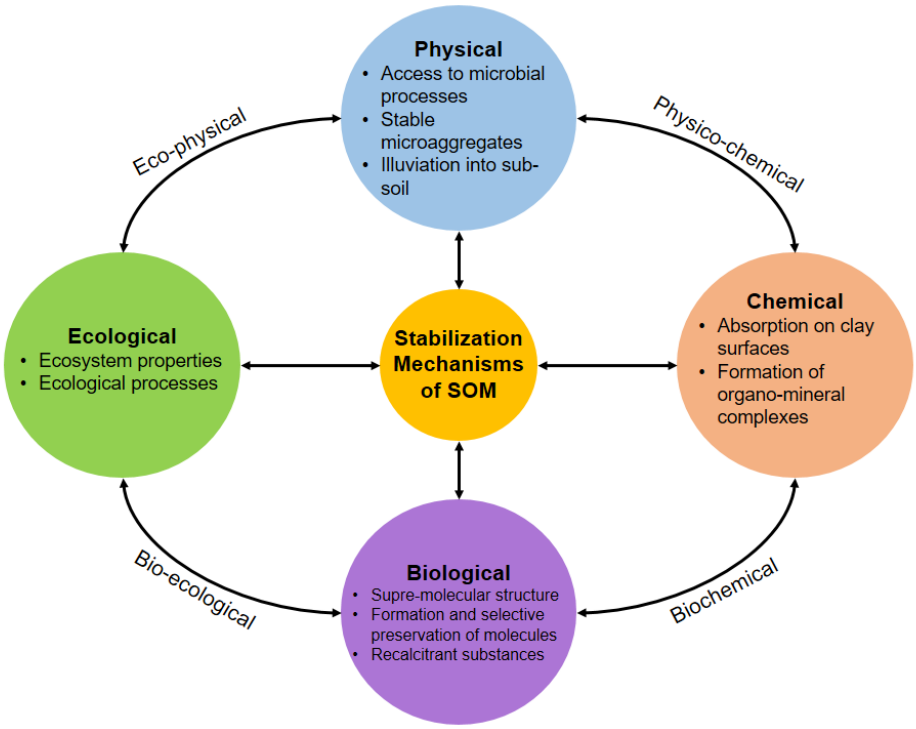


Figure 1.3 A wide range of mechanisms of stabilization of soil organic matter (adapted from Lal, 2017).

one often indicated as rSOC, obtained with strong NaOCl oxidation, as described in Mikutta et al. (2006).

Several studies have link empirically diverse carbon fractions to the theoretical components of SOC in more complex turnover models (Lehmann and Kleber, 2015). However, an empirical link, between the distribution of carbon in fractions and pools, does not necessarily mean a functional link, with model pools with a similar turnover or respond to changes in a similar way (Poeplau et al., 2013). Only few SOC models have been validated by measurable fractions, which in turn originate from different fractionation techniques (Zimmermann et al., 2007). Typical fractionation techniques include either physical or chemical (wet) separation of SOM or a combination of both (Poeplau et al., 2018). Physical fractionation are based on the premise that the association of soil particles and their spatial arrangement play a key role in SOC dynamics, because bioaccessibility is a prerequisite for decomposition (Christensen, 2001). Accordingly, the fractions are isolated by different degrees of disaggregation, dispersion, density fractionation and particle size separation (Poeplau et al., 2018). Aggregate fractionation is based on the separation of free SOM and protected SOM that is occluded in secondary organo–mineral assemblages of different sizes (Mikutta and Kaiser, 2011). Similar to physical fractionation, density one is applied to isolate SOM that is not firmly associated with soil minerals (light fraction) from organo–mineral complexes (heavy fraction) (von Lützow et al.,

2007). Lighter fractions consist mostly of pieces of plant residues, as particulate SOM (POM), either free (fPOM) or occluded in aggregates (oPOM) (Christensen, 1992).

Instead, chemical fractionation reflects the view that stability and turnover of each fraction is determined by its chemical composition. Hence, wet chemical procedures include fractionation of SOM according to solubility, hydrolysability, and resistance to oxidation or by thermal destruction (von Lützow et al., 2007). In particular, extraction of humic substances is popular because they generally extract large quantities of humic material in most soils considered to be refractory, thus belonging to the passive SOM pool (Hayes and Clapp, 2001). The combination of both procedures typically employs physical fractionation by using density or particle size fractionation to separate labile SOM fractions from stable SOM bound to minerals in the first step. In general, there are already differences in terms of accuracy of measured total SOC content, between dry and wet combustion. Chatterjee et al. (2009) have done an overview of these techniques. Wet combustion procedure involves oxidizing SOM to $CO_2$ with a solution containing potassium dichromate ($K_2Cr_2O_7$), but this reaction is insufficient to oxidize carbonaceous matter. As reported in more popular wet combustion method by Walkley and Black (1934), the incomplete digestion of organic materials assuming 76% recovery, is replaced by using the correction factor of about 1.30, as suggested by Santi et al. (2006) for Italian calcareous soils. Moreover, recovery values tends to increase when passing from silt-loam and loam to sandy soils, and decrease from deepest soil horizons to superficial ones (De Vos et al., 2007).

On the contrary, dry combustion methods lead to an overestimation of organic materials in the soil, compared to wet combustion, because of involving of oxidizing of SOC at a high temperature. From loss-on-ignition (LOI) procedure, the SOM is assessed by measuring the weight loss from a dry soil sample (oven-dried at 105°C) after high temperature ignition of the carbonaceous compounds in a muffle furnace. LOI to SOC conversion generally requires a correction factor of about 0.58, based on assumption that SOM comprises 58% of SOC. However, several researches (Hoogsteen et al., 2015; Jensen et al., 2018) indicated the inaccuracy of conversion, due to different soil types and different concentration of inorganic carbon and clay minerals in the studied soils (Ben-Dor and Banin, 1989). From these results given by a considerable improving of conversion from LOI to SOC, LOI parameter should be also corrected on the basis of clay content, to avoid over-estimating the SOM content by correcting for structural water loss.

The other dry combustion method involves the automated analyzer for determining total C content, via oxidation of both inorganic and organic carbon fractions to $CO_2$, treating the soil

samples at 950–1150°C. As rightly suggested by Chatterjee et al. (2009), measurements of C pools using automated analyzers are becoming standard procedure, due to several positive benefits produced, including the high accuracy of measurements, no loss of soil C during combustion and potential for simultaneously measuring nitrogen and sulfur. Dry combustion with auto analyzers have higher precision than wet combustion or LOI, due to possibility to detect even the inorganic fractions of carbon and those more recalcitrant of organic part, which instead remain with the others procedures, or are only in part degraded, contributing to misleading measurements. Table 1.1 shows the specifications of the different measurement techniques of the SOC.

Table 1.1
Comparison of methods used for determination of total C in soils (from Nelson and Sommers, 1996).

| Method | Principles | $CO_2$ determination | Advantages | Disadvantages |
|---|---|---|---|---|
| Dry combustion (resistance furnace) | Sample is mixed with CuO and heated to 1,000°C in a stream of $O_2$ to convert all C in sample to $CO_2$. | Gravimetric Titrimetric | Reference method widely used in other disciplines. Variable sample size. | Time-consuming. Leakfree $O_2$ sweep train is required. Slow release of $CO_2$ from alkaline earth carbonates. |
| Dry combustion (induction furnace) | Sample is mixed with Fe or accelerators and rapidly heated to >l,650°C in a stream of $O_2$ to convert all C in sample to $CO_2$. | Gravimetric Titrimetric | Rapid combustion. High temperature ensures conversion of C to $CO_2$. | Leakfree $O_2$ sweep train required. Induction furnace is expensive. |
| Dry combustions (automated methods) | Sample is mixed with catalysts or accelerators and heated with resistance or induction furnaces in a stream of $O_2$ to convert all C in sample to $CO_2$. | Gas chromatography (thermal conductivity or flame ionization detectors) | Rapid and simple. Good precision. | Expensive equipment. Slow release of $CO_2$ from alkaline earth carbonates with resistance furnace. |
| Wet combustion (combustion train) | Sample is heated with $K_2Cr_2O_7$-$H_2SO_4$-$H_3PO_4$ mixture in a $CO_2$-free air stream to convert all C in sample to $CO_2$. | Gravimetric Titrimetric | Equipment readily available. Good accuracy. Easily adapted to analysis of solutions. Titrimetric analysis of $CO_2$ less subject to operator error. | Time-consuming. Gravimetric determination of $CO_2$ requires careful analytical techniques. Titrimetric determination of $CO_2$ less precise. |
| Wet combustion (Van Slyke-Neil apparatus) | Sample is heated with $K_2Cr_2O_7$-$H_2SO_4$-$H_3PO_4$ mixture in a combustion tube attached to a Van Slyke-Neil apparatus to a convert all C in sample to $CO_2$. | Manometric | | Great skill needed to operate equipment. Equipment somewhat expensive and easily damaged. |

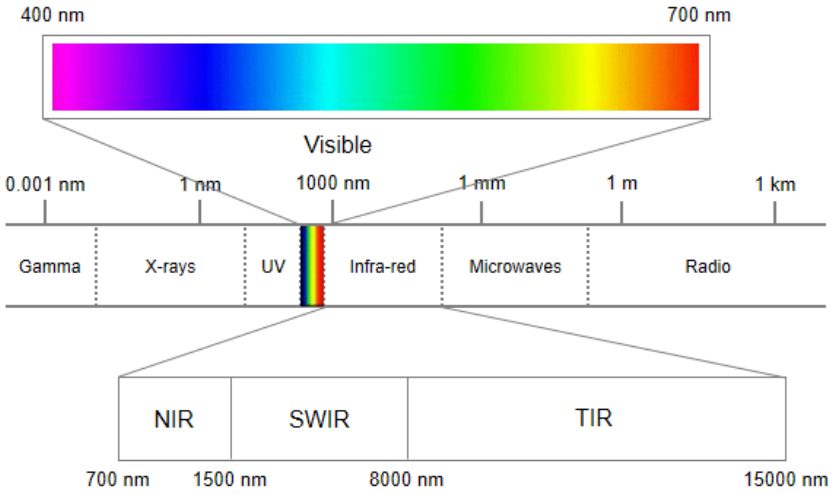


Figure 1.4 Electromagnetic spectrum in which the visible (Vis), near infrared (NIR), medium infrared (short-wavelength IR as SWIR) and thermal infrared (TIR) regions are highlighted.

Although soil sampling in the field and wet and dry combustion can be considered as the standard methods for assessment of C organic fractions of the soil, the whole process is expensive, time consuming and labor intensive. New techniques of soil carbon evaluation (predictive models) are based on remote sensing and spectroscopic measurements (Ben-Dor et al., 1999), even in the field, including near infrared (NIR) spectroscopy (Figure 1.4). The integration between the chemical laboratory analysis and spectral techniques (chemometrics) can contribute to the increasing of sustainability of measurements of organic carbon fractions of the soil, which continuously require deepest knowledge to comprehend well the role of the soil in mitigating climate changes.

## 1.2 Large scale and fast techniques for the assessment of soil organic carbon: proximal and remote sensing

Infrared reflectance spectroscopy is a rapid technique for measuring soil C based on the diffusely reflected radiation of irradiated soil (McCarty et al., 2002). In particular, the NIR spectral domain uses a quantitative determination of components of complex organic compounds, based on the absorption of the C-H, N-H, and O-H groups. Assessment and prediction of SOC using reflectance spectroscopy have been undertaken for several years, since Viscarra Rossel et al. (2006) and Torrent and Barrón (2008) laid the first solid foundations for the application of spectroscopy to the study of soil properties.

To give an overview, NIR spectroscopy can be considered as a proximal soil sensing (PSS) technique, which in turn is included among the remote sensing (RS) techniques. RS constitutes the set of techniques, tools and interpretative means that allow obtaining both qualitative and quantitative information on objects placed at distance. RS techniques are based on the collection of electromagnetic energy measured by special sensors. The principle on which remote sensing is based is the ability to differentiate as many elements as possible in an image at different territorial scales, thanks to their different radiometric behavior. In particular, the vegetation reflects light especially in the green and infrared bands, little in the blue and red bands; the behavior then differs, depending on the type of vegetation, the density, the phenological and phytosanitary state, the moisture content, etc. The reflectance of bare soils, on the other hand, depends on their physical (texture, structure, moisture content, etc.) and chemical (organic matter content, etc.) composition. The surface color is also a useful indicator of the type of soil and its properties (Gomarasca, 2016).

PSS is a remote sensing technique based on a series of measurement technologies in which the sensor is in direct contact or at a short distance from the soil surface. The advantage of these technologies, if applied for soil analysis, is the possibility of obtaining a large number of data quickly and at relatively low costs. The information is even used for the construction of high-detail pedological maps aimed at precision agriculture, as suggested by Robinson et al. (2008). What makes it possible to apply spectroscopy to soil analyzes is the presence of spectra-active compounds, which are called chromophores. A chromophore is a compound in which an atom or a group of atoms of a molecular entity may be present, responsible for a vibration following an electronic transaction. In the soil, there is an enormous amount of chromophores, both physical and chemical, likely responsible for the shape and nature of a spectral signature

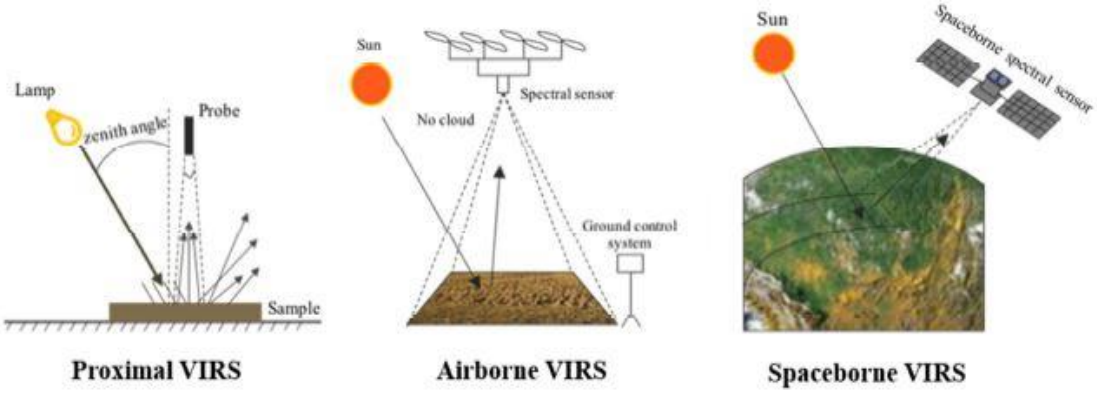


Figure 1.5 Schematic diagrams of proximal sensing, airborne imaging and a spaceborne spectrometer (VIRS, Visible and Infrared Spectroscopy) (Jia et al., 2021).

(Miller, 2001). These frequencies may be due to relatively mild vibrations of the C-N, N-H and O-H functional groups containing hydrogen, or to stronger bonds between the C-O, C-N, N-O and C-C atoms, present in the organic materials of the soil, as well as for the Al-O, Fe-O and Si-O bonds present, instead, in minerals (Janik et al., 1998; Coates, 2000; Du and Zhou, 2009).Overtones and combination bands in Vis-NIR due to organic matter result from the stretching and bending of N-H, C-H, and C-O groups (Ben-Dor et al., 1999). Bands around 1100, 1600, 1700 to1800, 2000, and 2200 to 2400 nm have been identified as being particularly important for SOC calibration (Ben-Dor and Banin, 1995; Clark, 1999; Stenberg, 2010).

However, spectroscopy in the Vis-NIR is largely nonspecific due to the overlapping absorption of soil constituents. Scatter effects caused by soil structure or specific constituents such as quartz compound this characteristic lack of specificity. Thus, as suggested by Stenberg et al. (2010), the analysis of soil diffuse reflectance spectra require the use of multivariate calibrations. When developing multivariate calibrations, careful selection of the calibration and validation samples is important. The size and distribution of the calibration data set has to be representative not only of the variation in the soil property being considered, but also of the variation in the spectra, according to Brown et al. (2005). In the last twenty years, reflectance spectroscopy and in particular Vis-NIR were widely used for the assessment of the SOC content in the soils. The need to pretreat the spectra through chemometric techniques was immediately clarified, and different approaches followed one another over the years. The most common calibration methods for soil applications are based on linear regressions, namely principal component regression (PCR) and partial least squares regression (PLSR). Several authors (Chang et al., 2005; Madari et al., 2005; Zornoza et al., 2008; Colombo et al., 2014; Mammadov et al., 2020) have used this regression technique achieving excellent results for prediction of SOC. In these cases, performances of prediction (measured vs. predicted) ranged between 0.73 and 0.94 in terms of values of $R^2$, suggesting that NIR was able to predict with reasonable accuracy total C, organic C and inorganic C. In particular, Zornoza et al. (2008) had tested NIR region with PLS approach even for prediction of biological and biochemical soil properties (microbial biomass carbon, basal soil respiration and enzymatic activity), achieving highly accurate results.

Others authors (Viscarra Rossel et al., 2006; Janik et al., 2007; Igne et al., 2010; Linsler et al., 2017) compared the performances of calibration for prediction of SOC in NIR and mid-IR (MIR) spectral ranges, indicating that in general MIR region allows to obtain more appreciable results than NIR. On the contrary, Rabenarivo et al. (2013) demonstrated that NIR provided more accurate predictions than MIR whatever the calibration procedure, a result related to the

abundance of some minerals in the studied soils and, as a consequence, the possible masking of organic matter peaks by the mineral component peaks in the MIR spectra, as they suggested.

In most recent years, the chemometric approach is changing, and the use of data mining techniques such as neural networks (NN), multivariate adaptive regression splines (MARS), boosted regression trees (BT) and random forest (RF) is increasing. Several authors have demonstrate that changing in regression model can be affect the accuracy of SOC prediction. In particular, support vector machine (SVM) regression outperforms PLS, as suggested by Lucà et al. (2017) and Asgari et al. (2020). Besides, SVM differently from PLS, which is comprised in linear regression methods, reduce high dimensional information through compressing the training samples to a smaller subset of important support vectors, after that utilize them for further model development. Viscarra Rossel and Behrens (2010) also tested the using Vis-NIR spectral information to predict SOC content with different regression approaches. In according to their research, SVM again allowed to obtain the best results, with predictions competitive to PLS. The NN approach was suitable only with a rather small number of variables, due to its non-linear interaction in the data, and, in agreement with Vasquez et al. (2009), they found tree ensemble approaches (RF and BT) performed weakly in the spectral domains for SOC calibration and validation.

Moreover, also samples and data pre-treatments highly contribute to Vis-NIR predictions (Figure 1.6). First, as suggested by Stenberg et al. (2010), the sieving of soil allows to ensure a constant particle size effect on the spectra, a feature that clearly distinguishes bench spectroscopy from field spectroscopy or other remote sensing techniques (Christy, 2008). After acquisition, it seems also advisable to pre-treat the spectra, in order to eliminate any effect of noise and scattering. The most common preprocessing is represented by the absorbance transformation of raw spectra, obtained through log 1/R, where R is the reflectance of the soil samples at any wavelength. However, many have computed SOC predictions across several pre-treatment techniques. Igne et al. (2010), Buddenbaum and Steffens (2012) and Riefolo et al. (2019) found prediction of SOC to be more performing when soil spectra were pre-treated with derivative transformation. Derivative is indeed considered to enhance spectral features, generally with a great impact on the visible and beginning of the NIR region, because absorption features in this region are very broad (Kooistra et al., 2001). On the contrary, Zhang et al. (2020) proved that only Savitzky-Golay smoothing filter had the capability to ameliorate prediction of SOM in saline soils despite all the others pre-treatment techniques.

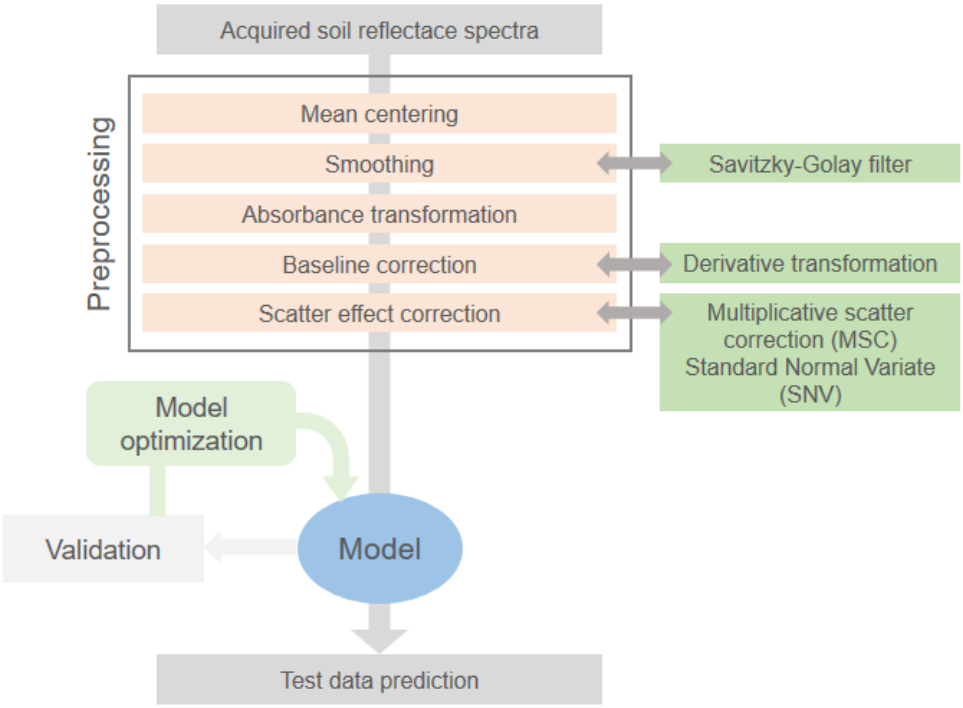


Figure 1.6 Spectra analysis for reflectance spectroscopy. The pretreatment steps remove corrupting effects that are not related to the sample. All these steps aim a robust prediction of the constructed model (adapted from Ryabchykov et al., 2018).

Summarizing, the combination between spectra pre-pretreatments and different regression models has allow to deepen the knowledge of organic carbon contents involving Vis-NIR soil characteristics. Application of Vis-NIR spectroscopy has also proved to be very useful for the prediction of the different organic carbon fractions of the soil, from the more labile and active ones to the stable and recalcitrant ones. Sarkhot et al. (2011) but also Viscarra Rossel and Hicks (2015) have gain very good results in predicting DOC as water-soluble organic carbon ($R^2$ of about 0.77), and particulate and recalcitrant organic carbon (POC and ROC), demonstrating the need to have techniques that allow for a fast and precise scanning of soil carbon pools, even on a large scale.

### 1.2.1 Sentinel-2 as a powerful tool for validating soil organic carbon pools

The need to obtain information on the content of the various organic carbon pools in soils in an ever more detailed way and with sustainable approaches even for monitoring at large-scale, has led researchers in recent times to make use of spectroscopy techniques linked to remote sensing. As previously said, RS is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance, and these data can come from different sources. Particularly, optical satellite multispectral imagery

started to be used extensively in quantitative SOC characterization with the launch of the first satellites in the 1980s, according to Frazier and Cheng (1989). Since, the applications based on hyperspectral data became popular, and have characterized the most recent studies.

As suggested by Angelopoulou et al. (2019), the functionality of the visible-near infrared and shortwave infrared Vis-NIR-SWIR sensors used for RS applications is based on the interaction principles between energy and matter, in the same way as for bench Vis-NIR spectroscopy. It means that correlating the spectral signatures with soil properties requires the use of multivariate statistical methods and chemometrics. The most common approach is the use of partial least squares (PLS) regression, which describes linear relationships between the variables, but machine learning algorithms (i.e. support vector machine, SVM) are increasingly used for the correlation process (Liakos et al., 2018). Differently, from bench Vis-NIR approach, RS techniques and in general spaceborne spectral imagery, are limited in use for soil observation, due to the required atmospheric, geometric and radiometric data corrections, simultaneous ground observations, the difficulty in finding large bare soil areas within a single image and obstacles related to vegetation cover (Castaldi et al., 2014).

RS as state-of-art method sees most of its application through the Sentinel-2 satellites use, characterized by high spatial, spectral and temporal resolution. Sentinel-2 carries an optical instrument payload that samples 13 spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution, useful for the retrieval and hence monitoring of SOC across the Vis-NIR spectral range. The applications of remote sensing data in SOC estimation have gain different results, and have been demonstrated to be suitable as auxiliary variables for mapping soil properties. Vaudour et al. (2019) have demonstrate that Sentinel-2 spectral bands are importantly affected from heterogeneity of data imagery acquisition on SOC prediction performances. Among these, soil roughness seemed to be the most relevant, in turn affected from the agricultural operations, crop residues and the main rainfall events, with the SOC prediction performances proved higher for spring dates. Dvorakova et al. (2020) also used PLS regression for SOC prediction across Sentinel-2 bands. They achieved discrete results for SOC prediction in terms of $R^2$, which reached its higher value (about 0.6) in correspondence of determined cellulose adsorption index (CAI), defining again the soil surface condition a relevant element for the accuracy of the estimate. Recently, Gholizadeh et al. (2018) proved the advantages of Sentinel-2 to derive high-quality information on variations in SOC comparing to airborne sensors, especially where SOC levels were relatively high. In that regard, they applied a simple SVM model to train prediction of SOC models. Others (Thaler et al., 2019; Zhou et al., 2021) have combine RS spectral information with the spatial prediction methods, as

ordinary or regression kriging, obtaining mapped soil properties including SOC with an acceptable precision.

As perfectly described in Angelopoulou et al. (2019), the main advantages of RS applications can be notice as non-destructive way to gather information about soil properties, the provided data cover large geographical areas, providing information about inaccessible areas for several attributes, and providing the means to reduce traditional and laborious soil sampling campaigns. However, the scarce information on these aspects calls for further investigate to reduce the negative impact of atmospheric absorptions and soil heterogeneity on SOC predictions, using satellite data. Anyway, Sentinel-2 is proving to be a powerful tool for soil fertility analysis, and in particular for the assessment and monitoring of soil organic carbon pools.

## 1.3 Aim of the thesis

Starting from the awareness that the soil plays a fundamental role in the mitigation of climate change through its ability to storage carbon in the form of organic matter, removing it from the atmosphere and thus decreasing the concentration of greenhouse gases, an ever greater analysis of the soil characteristics to implement the right management techniques that increase carbon sequestration is needed. In doing this, various methods for measuring and estimating soil carbon pools are now available. However, given the need to work on a large scale and in rather long periods, there is a need for reliable, rapid and inexpensive evaluation methodologies.

Placing in this broad framework of knowledge and insights, the aim of this thesis was to assess soil organic carbon pool of the studied organic farm, in order to know the impact of the organic and conservative soil management on the soil potential carbon sequestration. The steps through which the research took place were to: (i) analyze the soils from the point of view of their content of the different carbon fractions, both organic and inorganic, exploiting different available laboratory techniques and comparing them; (ii) apply Vis-NIR spectroscopy to the study of carbon pools of the soils to obtain a valid calibration model for their prediction through the use of different regression methods; and (iii) integrate predictive models by means of remote sensing techniques, in order to validate the models obtained and enhance the spatial variability of carbon pools from a field-scale to a farm-scale.

# Chapter 2
# Materials and Methods

## 2.1 Study area

The study area (Figure 2.1) includes territories of Melise farm, characterized by a mosaic land and located at Castel del Giudice municipality (Molise, south-central Italy) in the southern Apennines Mountains. This sector of the Apennine chain is characterized by a dominant hilly morphology and elevations ranging between 150 m (minimum height of main valley floors) and 1000 m a.s.l. (maximum height of some isolated peaks) (Aucelli et al., 2010). In particular, elevations of study area ranging between 750 and 1050 m a.s.l. (Figure 2.1).

From a geological point of view, all the territories of the northern part of Molise, on the border with the Abruzzo region, lie on calcareous and calcareous-marly Flysch, marl and, sometimes, varicolored clays. The characteristic of this area is that it is crossed by a series of ridges and valleys all parallel, in which the axes of the structure are directed in the Apennine sense (NW-SE). These are geological structures affected by folds, anticlines and synclines, sometimes interrupted by direct and reverse faults (fold-fault). The ridges coincide with the anticlinal structures or with the basal part of monoclinal structures, consisting of stratified

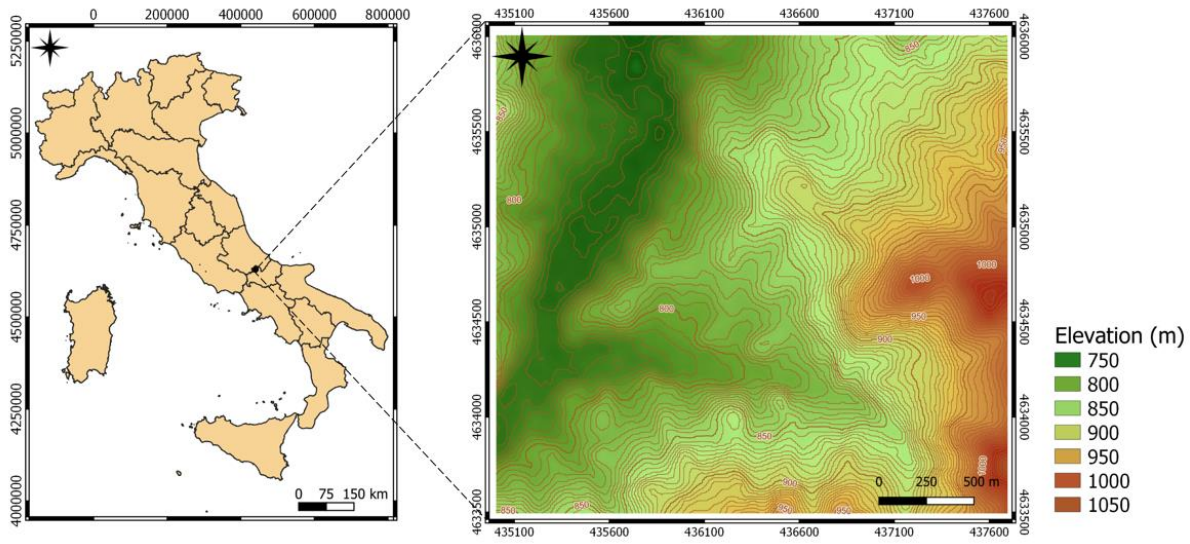


Figure 2.1 Geographical location and altimetric map of the entire area of Melise farm.

calcareous and calcareous-marl outcrops of the Miocene flyschoid complex. The valleys, on the other hand, coincide with the synclines, in which predominantly marly terms emerge, much more erodible than the calcareous ones, with classic V-shapes.

The morphology of this landscape reflects and is characterized by the different lithicity of the outcropping terms: the limestone, much more rigid than the marly formations, offer a lot of resistance to erosion, therefore, giving rise to ridges with steep slopes. On the contrary, the marls give rise to an intermediate landscape between that of limestone and that of clays, with wavy and wavy shapes separated by narrow valley incisions. The elevations, in correspondence of the crests, can reach up to 1000 m a.s.l. (Monte Mauro), while in the valleys there are altitudes of about 500 m at the top of the impluvium and 200 m in correspondence with the grafts with the rivers. The primary permeability of intact limestones and marly formations is low, while the secondary permeability increases due to billing. The present soils are from shallow to thin, well drained, with absent superficial and internal stonyness, with a fine texture, from very calcareous to highly calcareous.

### 2.1.1 Agronomic, pedological and climatic characterization of the area

The soil survey was aimed at a specific area of the municipality of Castel del Giudice, falling within the territories of the Melise farm. The company was founded in 2003, with the main objective of recovering the land in a semi-abandonment state of the municipality of Castel del Giudice, through the cultivation of apples and other species. Right from the start, Melise farm has chosen to undertake a method of conducting and cultivating orchards based on organic farming, for the inclusion of its products in an eco-sustainable food system, trying to eliminate or reduce the current sources of pollution. This choice is also aimed at trying to differentiate the agricultural entrepreneurship of that territory, based mainly on the creation of local brands and on self-consumption. Furthermore, in this farm, the principles of conservation agriculture (CA) are applied by 20 years. CA focus to maintain the crop residues on the soil surface combined with permanent zero-tillage approach, in order to reduce the excessive water disturbance on topsoil, minimize water soil erosion and sustain the accumulation of organic matter and nutrients.

The area of Melise farm that contains the orchards has an extension of about 40 hectares, with different varieties of apple trees cultivated, both cosmopolitan and native Molise varieties (Golden delicious, Gala, Fuji, Lasa, Sansa and Golden orange), including local variety with historical name: Zitella, Gelata and Limoncella (Figure 2.4). All the orchards, with 3.5x0.8 spacing, were planted between years 2000-2006, grafted on clonal rootstock M9, which induces

a low vigor to the plant and a consequent rapid fruit bearing. Every 20 plants a pollinator variety was placed, in order to guarantee the presence of at least 15-20% of pollinators, necessary in a monovarietal plant. From the point of view of the management of the orchards, the inter-row is characterized by presence of continuous cover crops with the subsequent release of plant residues obtained from the mowing of the lawn. The practice of continuous soil cover prevents, first, the compaction of the soil, making it, therefore, more porous, stable and permeable. The fertilizations consist of the addition of manure and other organic fertilizers, which are localized. As regards the water supply, in all the orchards there is a micro-irrigation system, which allows a significant water saving and avoids the development of weeds in the rows. The average water requirement of an apple orchard is around 500-600 $m^3$/ha of water per season, with a greater demand for water by the plants in the period of simultaneous growth of shoots and small fruits.

From a pedological sampling, it was possible to evaluate soils in studied area. Figure 2.2 shows details about soil profile. Because of presence of cambic horizon (Bw), distinguished for evidence of weathering and for a strong HCl effervescence, soil has been classified as *Calcaric Cambisols*, with supplementary qualifiers of *Loamic*, *Densic* and *Humic*, due to a great presence of clay and organic carbon contents to a depth more than 50 cm (IUSS WRB, 2015).

For a climatic characterization, Figure 2.3 shows the thermo-pluviometric diagram, referring to the years 2002-2014 and provided by Protezione Civile of Molise region, which characterizes climate of the areas close to Castel del Giudice, in which there are territories of the Melise farm.

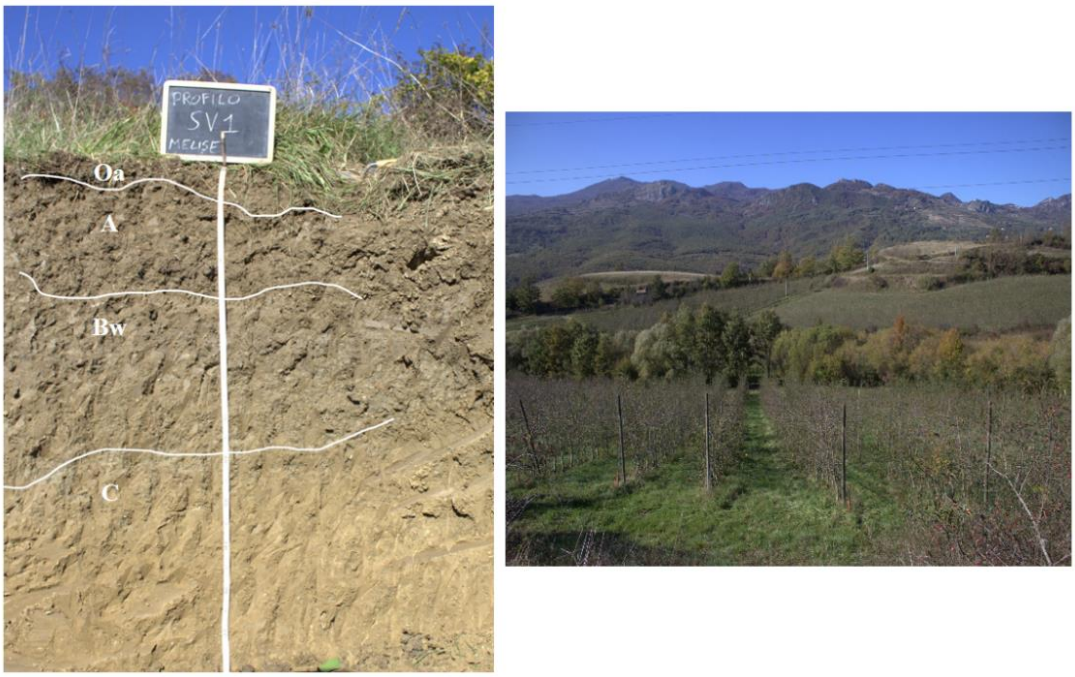


Figure 2.2 Characterization of soil profile sampled in studied area of Melise farm (436459 m E - 4634988 m N).

This area is characterized by a temperate climate, with mean annual rainfall and temperatures ranging respectively between 880 and 1550 mm and 5 and 23°C. Most of the annual precipitation occurs during the late autumn and spring months. Prolonged and extreme rainfall events occur periodically and are responsible for huge erosion related to both landslides and water erosion, in particular for area that are characterized by a value of slope greater than 20°, and in which soliflow phenomena are evident. Regarding mean annual temperatures, the study area is subject to severe frosts in the period between March and April, due to which the entire production can be compromised.

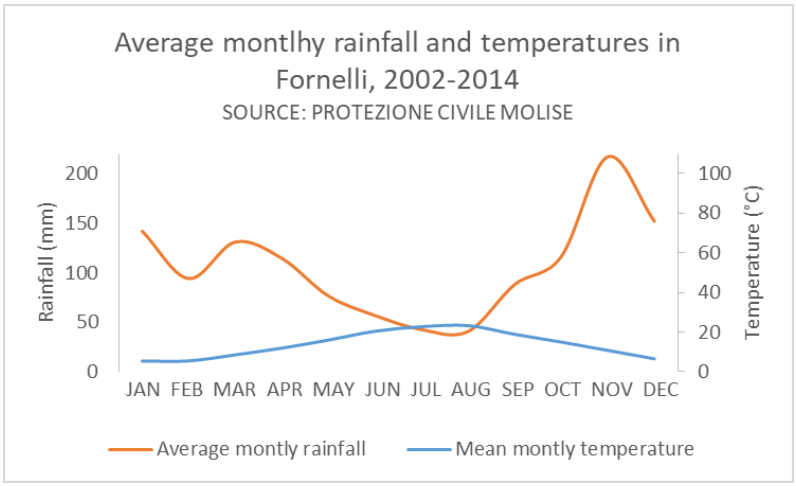


Figure 2.3 Climatic characterization of Fornelli area, municipality close to Castel del Giudice.

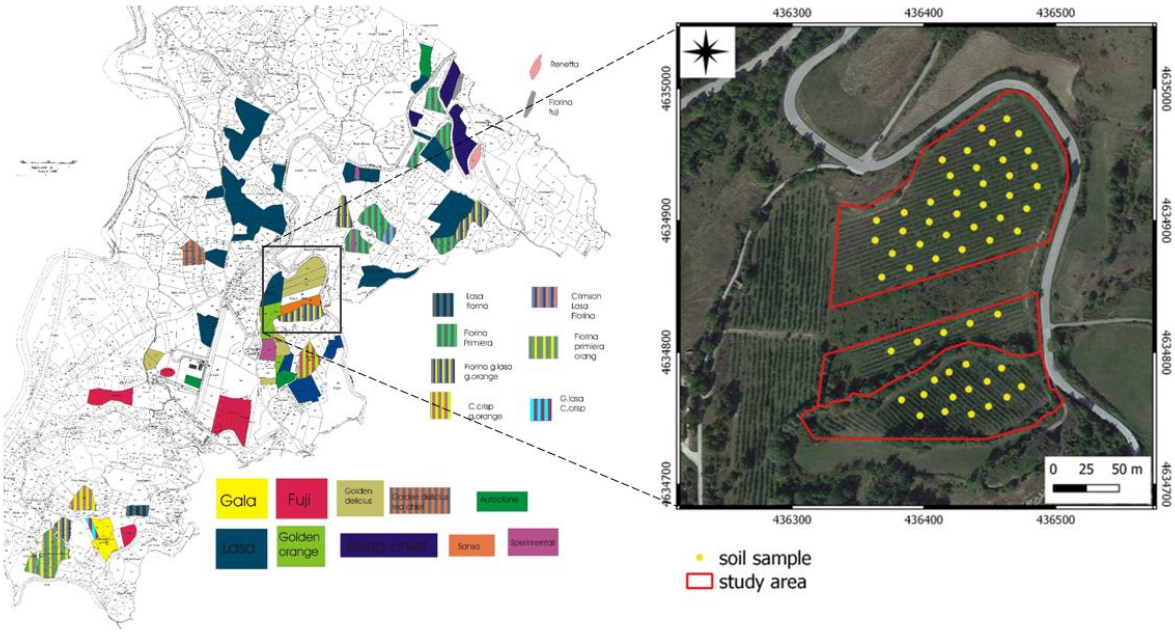


Figure 2.4 On the left, the cadastral map of Melise farm, with particulars in colors about cultivated apple varieties (striped colors are referred to re-grafts). On the right, the studied area with highlighted soil samples (yellow dots).

### 2.1.2 Experimental design

Soil sampling was carried out to highlight the spatial variability of the study area, minimizing the number of sampled points. In particular, it is a systematic topsoil sampling, at a distance of 20 meters and at a depth of 20 cm, according to the scheme shown in Figure 2.4, for a total of 56 georeferenced samples. Soil samples were then brought to the laboratoty and subposed to analyses discussed below.

## 2.2 Soil physico-chemical analysis

The investigated soils were analyzed for both physical and chemical properties. The analysis was performed on air-dried and 2 mm sieved samples.

The water content (WC) was determined by drying fresh soil at 105°C until a constant weight was reached, and the water holding capacity (WHC) was determined by the gravimetric method according to Aceves et al. (1994).

Soil pH was measured with a pH meter on aqueous extracts obtained by adding distilled water to the samples (2.5:1 = w/v), and electron conductivity (EC) determined on aqueous extracts of soil (1:5 = w/v) (MiPAF, 1999).

Soil particle size (soil textures) distributions were determined by pipette method (MiPAF, 1999). The method is based on the difference in sedimentation speed between small and large soil particles, applying the Law of Stokes. Soils were previously treated with hydrogen peroxide and sodium hexametaphosphate to remove respectively organic matter and aggregates. Classification was made according USDA system.

Bulk density (BD) was assayed on undisturbed soil cores of known volume after drying for 48 h at 105°C. Soil porosity was calculated according to Danielson and Sutherland (1986). For all soil sampling points, values of organic carbon Stock and saturated hydraulic conductivity (Ksat) were calculated. The first one is determined based on total organic carbon content, bulk density and depth of sampling (20 cm), in according to FAO (2017). Instead, Ksat coefficient, that represents the ability of soils to transmit water throughout the saturated zone, was estimated using a pedotransfer function (PTF), in terms of clay, silt and organic matter contents and bulk density, developed by Wösten et al. (1999) from the HYdraulic PRoperties of European Soils database (HYPRES).

Values of assimilable phosphorous (P) were determined via spectrophotometric in according to Olsen method (MiPAF, 1999), on soil solution aliquotes previously extracted with sodium bicarbonate solution.

Soils were also analyzed for cation exchange capacity (CEC), and with the application of this method, the extraction through barium chloride and triethanolamine solution led to quantification of exchangeable bases (Na, Ca, K and Mg).

At last, concentrations of assimilable Fe, Mn, Cu and Zn were carried out through extraction with DTPA solution as indicated by Lindsay and Norwell (1978) method. Both exchangeable bases and assimilable cations concentrations were recording by atomic absorption (AA) spectrometry (Thermo Scientific iCE$^{TM}$ 3300).

### 2.2.1 Carbon fractionation

For evaluation of the different fractions of soil organic matter (SOM), several analytical methods were applied (Figure 2.5). Firstly, total organic carbon ($TOC_{WB}$) content was carried out through wet-oxidation with potassium dichromate by Walkley-Black system (Walkley and Black, 1934). Total carbon ($TC_{CNS}$) and nitrogen ($TN_{CNS}$) contents was also determined via gas chromatography (Thermo Finnigan, CNS Analyzer) (Colombo and Miano, 2015), and with the same procedure, organic carbon ($OC_{CNS}$) was measured on soil samples previously treated with hydrogen chloride (HCl 10%) to exclude carbonates. Successively, C/N ratio was calculated such as the inorganic C fraction ($IC_{CNS}$), through the difference between $TC_{CNS}$ and $OC_{CNS}$.

To deepen the knowledge on the organic and inorganic carbon content of the investigated soils, loss-on-ignition (LOI) procedure was also performed as well as indicated by Nelson and Sommers (1996). Five grams of air-dry soil were added to previously ignited and weighed porcelain crucibles, dried at 105°C for 24 hours in a ventilated oven, cooled in a desiccator and weighed again. The crucibles were ignited at 500°C for 4 hours in a muffle furnace (Muffle Furnace, Thermo Fisher Scientific, Waltham, MA, USA). After ignition, the crucibles were cooled in a desiccator and weighed. The LOI was calculated as the difference between the oven-dry weight before and after ignition and related to oven-dry soil. This method estimates soil organic matter based on gravimetric weight change associated with high temperature oxidation of organic matter, although the application of this method generates an overestimation error, due both to the presence of carbonates and gypsum, and to the loss of water of clay minerals (Ben-Dor and Banin, 1989). Loss-on-ignition (LOI), however, remains a widely used method for assessing SOM in agricultural and forest soils and being converted to SOC either by a fixed conversion factor or by regression analyses (Jensen et al., 2018). After, in order to complete

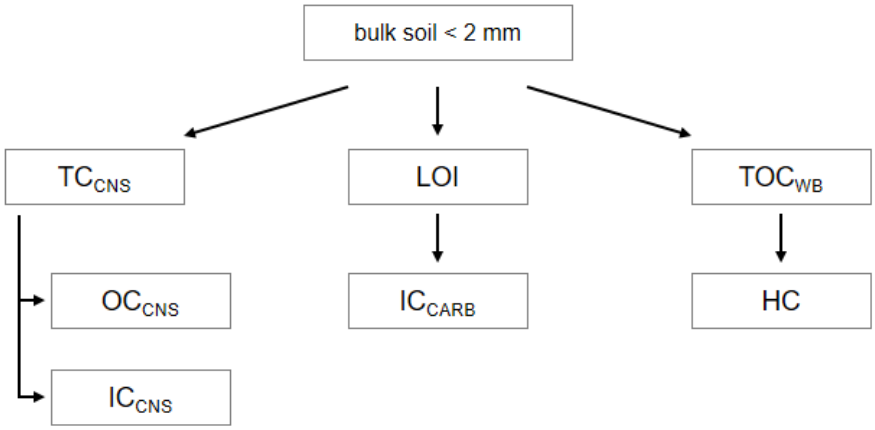


Figure 2.5 Schematic description of the soil carbon fractionation procedure. Total carbon fraction ($TC_{CNS}$) that can be distinguished in the organic $OC_{CNS}$ and inorganic $IC_{CNS}$ fractions. A second carbon procedure was the loss-on-ignition (LOI), and the inorganic carbon content ($IC_{CARB}$) was determined as $CaCO_3$. Wet-oxidation procedures with potassium dichromate were applied to obtain total organic C fraction ($TOC_{WB}$) and the soluble organic carbon (HC), previously extracted with hot water.

analysis on soil organic carbon pools and to have data on easy degradable or available fraction of organic carbon present in soils, hot water extractable carbon (HC) was measured by using extraction procedures from Sarkhot et al. (2011). Briefly, soil water mixture (1:5 respectively) was heated at 80°C for 16 h, the soil suspension was centrifuged for 3 min at 2000 rpm and filtered through a 0.22 μm GV membrane filter (Durapore). The filtrate was air-dried, weighted, analyzed by spectroscopy technique and, only after this, it was examined through wet-oxidation with potassium dichromate according to Springer-Klee method (Springer and Klee, 1954), in order to know the quantity of organic carbon in the extracted fraction.

Total calcium carbonate ($IC_{CARB}$) as inorganic carbon content was determined by acid solubilization with hydrogen chloride (1:1 v/v), by measuring the amount of $CO_2$ produced by the reaction. The instrument that was used for this purpose is the Dietrich-Fruehling calcimeter, which performs the measurement by gas-volumetric method with internal calibration using pure $CaCO_3$. Instead, active pools of calcium carbonate was carried out through extraction with ammonium oxalate and titration with potassium permanganate solution (MiPAF, 1999).

## 2.3 Soil biological analysis

The investigated soils were analyzed for microbial biomass as well as microbial respiration. The analysis were performed in triplicates on air-dried soil samples pre-conditioned at field capacity.

Microbial biomass carbon ($C_{MIC}$) was evaluated by the method of substrate-induced respiration (SIR) according to Degens et al. (2001), while microbial potential respiration (Resp) according to Froment (1972). The $CO_2$ evolution from the samples was measured by NaOH absorption followed by two-phase titration with HCl, after incubation at 25 °C in tight containers for 5 and 10 days, respectively, to evaluate $C_{MIC}$ and Resp.

In order to make comparable the soil samples collected at different sites and water contents, all data were expressed per unit of soil dry weight. The results obtained by the biological analysis were then used to calculate two indices: the metabolic quotient ($qCO_2$), which represents the degree of microbial biomass activity, and the coefficient of endogenous mineralization (CEM), that is the rate of organic carbon mineralization (Anderson and Domsch, 1993). The $qCO_2$ was calculated as ratio between Resp and $C_{MIC}$, whereas the CEM was calculated as ratio between Resp and $TOC_{WB}$.

## 2.4 Statistical analysis of soil properties

The measured soil properties were statistically described in terms of minimum, maximum, mean, median, standard deviation (SD) and variance. The variance is a measure of the spread of observations around a mean value. It is calculated as the sum of squares of the individual observations and the mean, divided by the degrees of freedom (DOF) associated with the observations. Additionally, Skewness and Kurtosis indexes were determined. The Skewness of a distribution is a measure of its asymmetry and is referred to as the third central moment of the distribution. The Kurtosis describes the extent of the degree of flatness of the center of a distribution.

The correlation matrix between soil physico-chemical and biological properties was also computed, based on Pearson's correlation coefficient. The statistical significance of results was measured according to values of *p*-level. The descriptive statistics was calculated in Statistica 10 software (StatSoft Inc., Tulsa, OK, USA).

With the aim of reveal the hidden structure within dataset and provide a visual representation of the relationships among samples and variables (e.g. sample groupings, similarities or differences), principal component analysis (PCA) was carried out by Unscrambler software (Camo, Inc.). Variables were previously standardized and all were used, including the highly correlated pairs, in order to studying the variables' interaction.

## 2.5 Soil spectra analysis

Soil samples (air-dried and 2 mm sieved) were analyzed by Vis-NIR diffuse reflectance spectroscopy (DRS). The reflectance curves were obtained using a Jasco spectrophotometer, model V-570 Vis-NIR, equipped with double beam and double monochromator (Jasco Inc. Easton, MD), and with an integration sphere, having a 73 mm diameter and coated with barium sulphate $BaSO_4$ (Jasco ISN 470 integrating sphere system). The geometry of the integration sphere (Figure 2.6), through the presence of two openings, allows the simultaneous measurement of the sample and the blank, the barium sulphate, which has a reflectance equal to 0.988 (Torrent and Barrón, 2008). Only the scattered light is analyzed by the sensor, the specular one is excluded by dispersion and reflection outside the sphere.

The diffuse reflectance spectrum is significantly influenced by the sample pretreatment mode. Particle size is the factor that most affects reflectance. An example is provided by the evolution of the color of the samples, sometimes even consistent, after being subjected to heavy grinding. According to Torrent and Barrón (2008), the best results were obtained with rather small particle sizes, by finely grinding the samples down to diameter sizes of less than 10 μm; for sandy samples, on the other hand, this treatment can cause a drastic change in color, due to the different colors of the primary minerals that compose them.

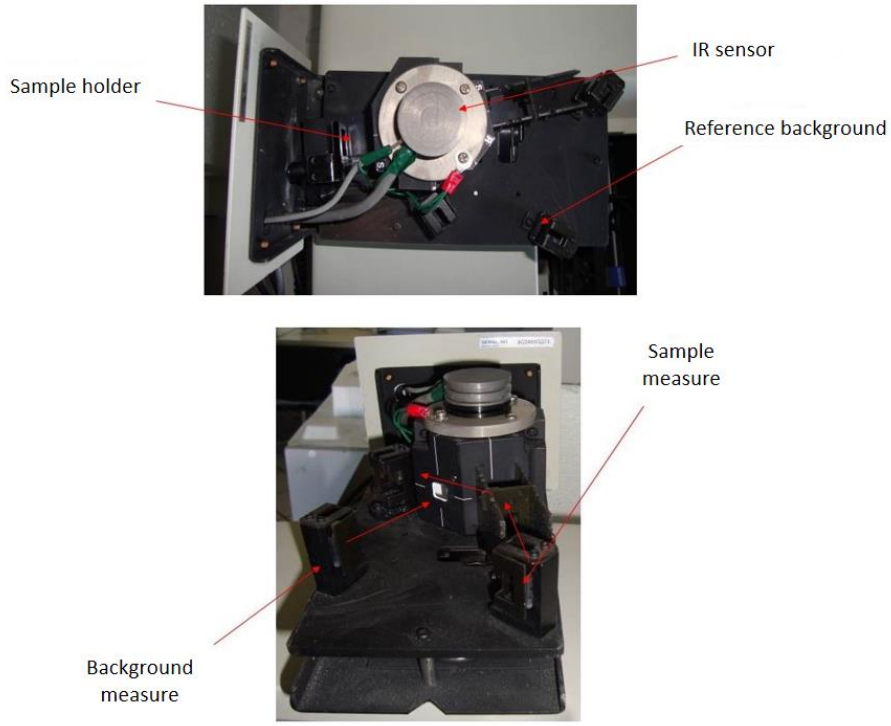


Figure 2.6 Geometry of the integration sphere in the spectrophotometer. The blank and sample supports are highlighted, as is the path of the incident light beam.

The intensity of the grinding also affects the spectral results. As can be seen from studies reported in the literature (Kortüm, 1969; Fernandez and Schulze, 1987), excellent results have been obtained in the conservation of the spectral characteristics by finely and delicately grinding the soil samples in an agate mortar for about 10 minutes. In this way, the particle size of the minerals decreases uniformly without favoring the scattering process (Barrón and Torrent, 1986).

Soil samples were then ground in an agate mortar in order to homogenize and uniform the particles. A small amount of soil was gently inserted into special slides, which have a rectangular cavity, opaque after treatment with corundum powder. The slide is subsequently placed vertically on the integration sphere. Spectra were acquired in the spectral range of 200-2500 nm, with a scan resolution of 2 nm and scanning speed of 1000 nm/min, for a total of about 1150 points. Each Vis-NIR spectrum was recorded as percent reflectance (R%). The analysis of the spectra were carried out using the software supplied with the spectrophotometer, Spectra manager (JASCO Corp.).

### 2.5.1 Data preparation and preprocessing transformations

The preliminary stage of spectroscopy analysis is generally preprocessing to evaluate and possibly increase data quality (see Figure 1.6). Spectral preprocessing method consists of different mathematical techniques for correcting light scattering in reflectance measurements and data improvement before employing data in calibration models (Gholizadeh et al., 2015).

Here, five different pre-processing methods were applied in order to compare the obtained results of multivariate calibration with the one obtained by raw reflectance data: 1) Savitzky-Golay smoothing algorithm combined with absorbance transformation (SG+Abs); 2) multiplicative scatter correction (MSC); 3) standard normal variate (SNV); 4) first derivative and 5) second derivative transformations.

The Savitzky-Golay algorithm (Savitzky and Golay, 1964) fits a polynomial to each successive curve segment, thus replacing the original values with more regular variations. It is a very useful method to effectively remove spectral noise spikes while keeping chemical information. In this case, smoothing at 11 smoothing points (5 points left and 5 points right) and second polynomial degree were chosen.

Transformation from reflectance to absorbance units is given by $\log_{10}$[1/reflectance], which can be related to concentration by Beer's law. Conversion into absorbance units is important to make more intense the visible range of the adsorption spectra that are generally well correlated with measured soil properties (Riefolo et al., 2019).

Multiplicative scatter correction (MSC) is a transformation method used to compensate for pure additive and multiplicative effects in spectral data. Elaborated by Geladi et al. (1985), MSC removes artifacts or imperfections (e.g. undesirable scatter effect) from the data matrix, simplifying the calibration model through reduction of the number of components needed and improving the linearity (Næs et al., 2002).

Standard normal variate (SNV) is a transformation applied to spectroscopic data to remove scatter effects by centering and scaling each individual spectrum (i.e. a sample-oriented standardization). Like MSC, the practical result of SNV is that it removes multiplicative interferences of scatter and particle size effects from spectral data. Spectra are centered on zero and varies roughly from -2 to +2. Both MSC and SNV methods are designed to reduce the (physical) variability between samples due to scatter (Rinnan et al., 2009).

Derivatives are applied to correct for baseline effects in spectra for the purpose of removing nonchemical effects and creating robust calibration models. Derivatives may also aid in resolving overlapped bands that can provide a better understanding of the data, emphasizing small spectral variations not evident in the raw data (Viscarra Rossel, 2008; Colombo et al., 2014; Rienzi et al., 2014). Particularly, the first derivative of a spectrum is a measure of the slope of the spectral curve at every point; instead, the second derivative is a measure of the change in the slope of the curve. It means that in addition to removing pure additive offset (effect of first derivative), second derivative removes both the baseline offset and slope from a spectrum. In this case, derivation techniques were combined with Savitzky-Golay smoothing in order not to reduce the signal-to-noise ratio in the corrected spectra too much.

Spectral data transformations and preprocessing were performed by Unscrambler software (Camo, Inc.). Figure 2.7 depicts the raw reflectance spectra and their preprocessing of all soil samples.

### 2.5.2 Prediction models development

The dataset (56 samples) for the entire spectra range (Vis-NIR) was randomly divided into calibration (70% of the samples: 39 samples) and validation (30%: 17 samples) datasets. Three different calibration models were performed over Vis-NIR spectra to develop models capable of predicting soil properties, including Partial Least-Squares Regression (PLSR), Support Vector Machine (SVM) and Boosted Tree (BT). Cross-validation was performed to determine the number of latent variables that minimizes the predicted residual sum of squares, as indicated by Riefolo et al. (2019).

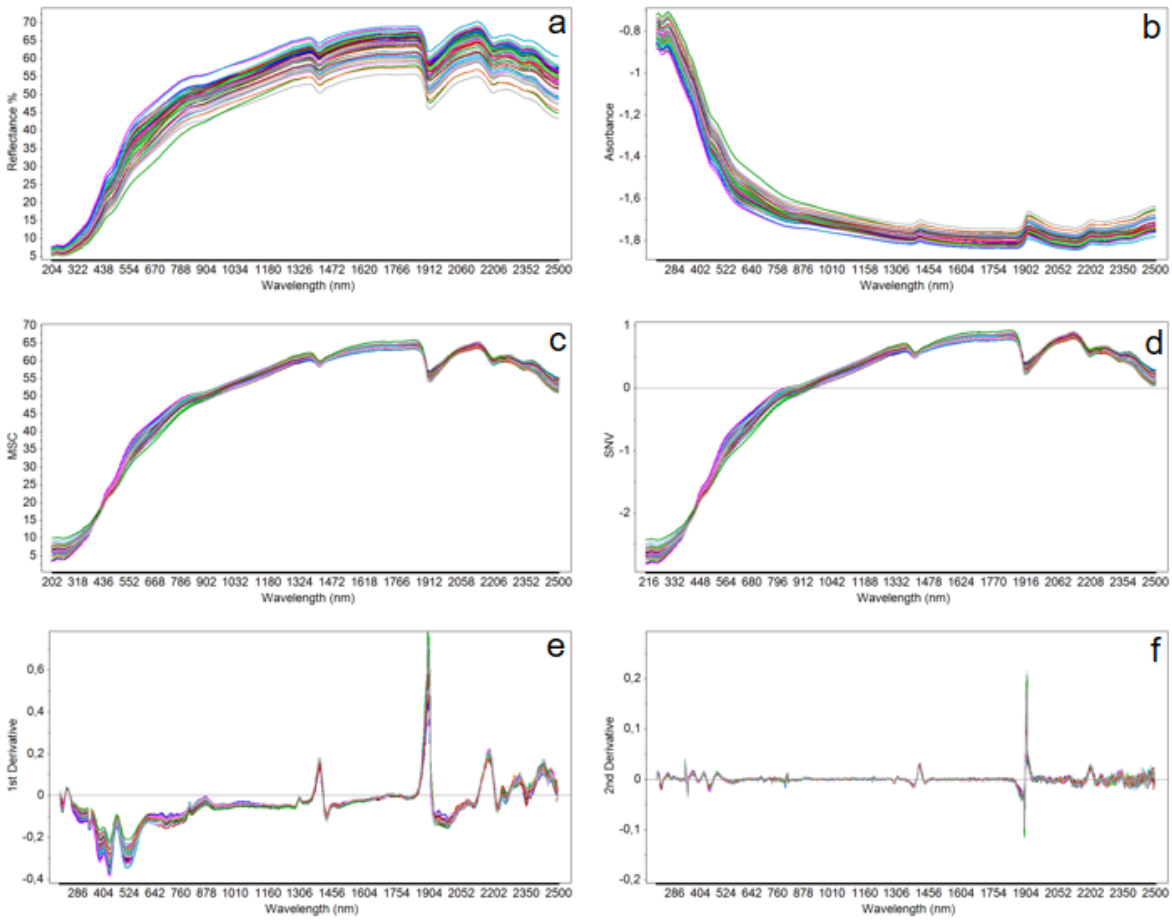


Figure 2.7 Spectra preprocessing of all soil samples: a) raw reflectance spectra; b) Savitzky-Golay smoothing and transformation into absorbance units of raw spectra; c) multiplicative scatter correction (MSC) on raw spectra; d) standard normal variate (SNV) on raw spectra; e) first derivative and f) second derivative transformations on raw spectra.

Partial least square regression (PLSR), developed by Wold et al. (1983), is a latent-variable-based method for the linear modelling of the relationship between a set of response variables Y and a set of predictor variables X. The PLSR algorithm selects successive orthogonal factors that maximize the covariance between predictors (X spectra) and response variables (Y laboratory data). It is the most commonly used method for designing predictive soil attributes models from diffuse reflectance spectra (Asgari et al., 2020).

Support vector machine (SVM) is a kernel-based learning method from statistical learning theory, originally proposed by Vapnik (2000). Differently from PLSR, SVM regression acts as a nonlinear regression technique, where the complexity of the nonlinear regression curve is controlled via the number of support vectors. Hence, it is a nonparametric technique.

Boosted tree (BT), as described in Friedman et al. (2000), is a full featured implementation of the stochastic gradient boosting method. The general idea of this type of regression is to compute a sequence of simple trees, where each successive tree is built for the prediction

residuals of the preceding tree. The method of gradient boosting, as well as SVM, represents a very general and powerful machine learning algorithm.

PLSR, SVM and BT regressions were performed using Statistica 10 software (StatSoft Inc., Tulsa, OK, USA). The performance of three prediction models was evaluated by means of three statistics: 1) the coefficient of determination in prediction ($R^2$) (Eq. 1), 2) root mean square error (RMSE) (Eq. 2), and 3) residual prediction deviation (RPD) (Eq. 3).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (1)$$

$$RMSE_{val} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

$$RPD_{val} = \frac{SD_{val}}{RMSE_{val}} \quad (3)$$

where y, ŷ and ȳ are the measured values, predicted values, and mean of measured values respectively, and n indicates the soil samples number. SD is the standard deviation of predicted values and $RMSE_{val}$ is the root mean standard error of validation dataset. As indicated in Di Iorio et al. (2019) and according to several authors (Chang et al., 2001; Dunn et al., 2002; Cozzolino and Moron, 2003), RPD values below 1.5 indicate a poor predictive model; in the range 1.5-2.0 they point to acceptable results, which may be not enough for exactly estimating the target variable (qualitative prediction); RPD values higher than 2.0 are considered excellent and indicate a prediction suitable for quantitative reliable estimations.

Variable Importance for Prediction (VIP) statistics was also used for the selection of relevant wavelengths. VIP statistics synthetizes the contribution of a variable to the model because it is a weighted sum of squares of X-score coefficients for the retained components, with the weights calculated from the amount of dependent variable (y) variance explained by each retained component (Wold et al., 2001). A threshold level of 3.0 was considered to identify the most important wavelengths representing overtones and bands of fundamental molecular vibrations, as done in Riefolo et al. (2019).

## 2.6 Sentinel-2 data collection

The satellite images used in this study include Sentinel-2 Multispectral Instrument (S2-MSI) missions, obtained from European Space Agency (ESA). Each Sentinel-2 mission is based on a constellation of two satellites (2A and 2B) to fulfil revisit and coverage requirements, providing robust datasets for Copernicus services. These missions carry a range of technologies, such as radar and multi-spectral imaging instruments for land, ocean and atmospheric monitoring. In particular, Sentinel-2 is a polar-orbiting, multispectral high-resolution imaging mission for land monitoring to provide imagery of vegetation, soil and water cover, inland waterways and coastal areas. Table 2.1 shows details about spectral bands of Sentinel-2.

The Sentinel-2 satellite images, for both Levels 1C and 2A, were downloaded from the Copernicus Open Access Hub platform, for ten different dates of the tile T33TVG (Table 2.2). The Level-1C product results from using a Digital Elevation Model (DEM) to project the image in cartographic geometry. Per-pixel radiometric measurements are provided in Top Of Atmosphere (TOA) reflectances along with the parameters to transform them into radiances. Instead, the Level-2A product provides Bottom Of Atmosphere (BOA) reflectance images derived from the associated Level-1C products, through atmospheric correction.

Table 2.1
Specifications of a Sentinel-2 Multispectral Instrument sensor, provided by Copernicus Open Access Hub platform.

| Sentinel-2 Bands | Sentinel-2A | | Sentinel-2B | | Spatial resolution (m) |
|---|---|---|---|---|---|
| | Central Wavelength (nm) | Bandwidth (nm) | Central Wavelength (nm) | Bandwidth (nm) | |
| Band 1: Coastal aerosol | 442.7 | 21 | 442.2 | 21 | 60 |
| Band 2: Blue | 492.4 | 66 | 492.1 | 66 | 10 |
| Band 3: Green | 559.8 | 36 | 559.0 | 36 | 10 |
| Band 4: Red | 664.6 | 31 | 664.9 | 31 | 10 |
| Band 5: Red-edge 1 | 704.1 | 15 | 703.8 | 16 | 20 |
| Band 6: Red-edge 2 | 740.5 | 15 | 739.1 | 15 | 20 |
| Band 7: Red-edge 3 | 782.8 | 20 | 779.7 | 20 | 20 |
| Band 8: NIR | 832.8 | 106 | 832.9 | 106 | 10 |
| Band 8A: narrow NIR | 864.7 | 21 | 864.0 | 22 | 20 |
| Band 9: Water vapour | 945.1 | 20 | 943.2 | 21 | 60 |
| Band 10: SWIR Cirrus | 1373.5 | 31 | 1376.9 | 30 | 60 |
| Band 11: SWIR | 1613.7 | 91 | 1610.4 | 94 | 20 |
| Band 12: SWIR | 2202.4 | 175 | 2185.7 | 185 | 20 |

Table 2.2
Main characteristics of the studied scenes (tile T33TVG): Sp, spring; Su, summer; A, autumn; W, winter.

| Season | Imaging Date | Time of Acquisition | Sun Zenith (°) | Sun Azimuth (°) | Cloud Cover of the entire tile (%) |
|---|---|---|---|---|---|
| Sp | 3 April 2018 | 09:58:07 | 39.5 | 153.4 | 49 |
| Sp | 23 April 2018 | 09:54:54 | 32.3 | 151.1 | 5.7 |
| Sp | 13 May 2018 | 09:57:05 | 26.7 | 147.3 | 9.9 |
| Sp | 2 June 2018 | 09:58:59 | 23.5 | 142.4 | 10.6 |
| Su | 5 July 2018 | 10:05:08 | 22.7 | 143.4 | 0.3 |
| Su | 11 August 2018 | 09:56:12 | 30.5 | 146.1 | 0.8 |
| A | 28 September 2018 | 10:08:05 | 45.1 | 164.5 | 0.2 |
| A | 25 October 2018 | 09:55:14 | 55.2 | 166 | 1.1 |
| A | 12 November 2018 | 10:08:58 | 60.3 | 170 | 6.2 |
| W | 29 December 2018 | 09:59:01 | 66.9 | 163.7 | 2.9 |

From reflectance images of Level-2A vegetation radiometric indexes were calculated for the entire study area. Normalized difference vegetation index (NDVI) (Eq. 4) (Rouse et al., 1973), provides a measurement of crop health. It describes the vigor level of the crop, since it allows to immediately recognizing the areas of the field that have development problems. Values can vary between -1 and 1. Similar to NDVI, normalized difference water index (NDWI) (Eq. 5) (Gao, 1996), describes the level of water stress of the crop, as it allows to immediately recognize the areas of the field that have water stress problems. Again, values can vary between -1 and 1.

$$NDVI = \frac{R_{NIR} - R_{red}}{R_{NIR} + R_{red}} \quad (4)$$

$$NDWI = \frac{R_{NIR} - R_{SWIR2}}{R_{NIR} + R_{SWIR2}} \quad (5)$$

where $R_{red}$, $R_{NIR}$ and $R_{SWIR2}$ are values of reflectance for bands 4, 8 and 12 respectively.

Prior to indexes calculation, the images were resampled to 10 m, which was the highest resolution amongst 10, 20, and 60 m. All satellite data pretreatment and processing were performed using the Sentinel Application Platform (SNAP) tool version 8.0.

Additionally, for the same tile and close to chosen plots of study area, pixels of bare soil were selected, in order to download spectra of soil and utilized them as an external validation dataset for prediction of unknown values of soil organic carbon. Firstly, the choice of pixels of bare soil was made by calculating the corresponding NDVI index. As done in Vaudour et al. (2019) and according to Loiseau et al. (2019), the threshold value of 0.27 was selected as the best trade-off for getting the largest bare soil area while reducing the effects of sparse vegetation on the field scale: pixels with an NDVI index greater than 0.27 were removed. After

downloading, each spectrum of satellite images was interpolated through natural cubic splines method to gain 2 nm resolution as for laboratory soil spectra, which were utilized as calibration dataset for construction of prediction models. Performances of prediction models were evaluated again with $R^2$, RMSE and RPD statistics.

The predictors used for calibration model (laboratory spectral data) and those used as external validation (Sentinel-2 spectra) were centered and scaled to have zero mean and variance 1 (Chilès and Delfiner, 2012), in order to place both on the same relative position to their variation in the data and to minimize dimensionality and differences.

Pin Manager Toolbox of SNAP tool made the downloading of Sentinel-2 spectra, interpolations and subsequently processing were made using function Predict of Unscrambler software (Camo, Inc.).

## 2.7 Geostatistical analysis

The geostatistical procedure is a linear regression method that allows you to interpolate a quantity in space, minimizing the mean square error. The interpolation is based on the assumption of the observations made in the neighboring points have a higher correlation than the observations placed in distant points (Laslett et al., 1987).

Spatial predictions of soil carbon pools were performed by using univariate ordinary kriging (OK) method (Webster and Oliver, 2007) in QGIS (v. 3.14) and ArcView GIS (v. 3.1) software. Kriging is an interpolation methods consists of geostatistical methods, which are based on statistical models that include autocorrelation, the statistical relationships among the measured points. Because of this, geostatistical techniques not only have the capability of producing a prediction surface but also provide some measure of the certainty or accuracy of the predictions, in this case by means of semivariogram. Cross-validation was used to derive the criterions to assess estimation quality. Performances of two applied models of geostatistical interpolation, spherical and exponential, were evaluated by means of Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) and root mean squared error (RMSE).

# Chapter 3
# Results and Discussion

## 3.1 Descriptive statistics of soil properties

### 3.1.1 Carbon fractions

The descriptive statistics of the investigated soil properties are reported in Table 3.1. Soil total carbon values ($TC_{CNS}$) which contained both organic and inorganic C fractions, ranged from 32.18 to 70.05 g $kg^{-1}$, with a mean value of 47.44 g $kg^{-1}$ and a symmetrical distribution, as indicated by Skewness and Kurtosis indexes. The application of CNS analyzer allowed to obtain the HCl treated fraction ($OC_{CNS}$) and the difference ($TC_{CNS}$ - $OC_{CNS}$) as inorganic carbon fraction ($IC_{CNS}$), both accounted for 52% and 48% of total carbon, respectively.

From the second procedure with wet oxidations analysis about different pool of organic carbon in the soil, it turned out that total organic carbon ($TOC_{WB}$) and available and active organic carbon (HC), accounted for 41% and 3% respectively of total carbon.

As possible to see in Table 3.1, the $TOC_{WB}$ values ranged from 1.93 to 49.12 g $kg^{-1}$, while the $OC_{CNS}$ values from 18.96 to 60.48 g $kg^{-1}$, $IC_{CNS}$ from 0 to 38.8 g $kg^{-1}$ and HC values ranged from 0.52 to 4.17 g $kg^{-1}$.

Figure 3.1a summarizes the main carbon fractions quantified with CNS analyzer, which can be distinguished in the organic $OC_{CNS}$ (52% of TC) and inorganic as $IC_{CNS}$ with the remnant part of about 48%.

The box plots of Figure 3.1b depict the distribution of soil organic carbon (SOC) among all the fractions observed under the thermal and wet chemical treatments. The largest part of SOC appeared in LOI. The fraction calculated as LOI - $TOC_{WB}$ seems to contain about 40% of inert carbon much higher than the carbon content in the carbonate ($IC_{CARB}$), which was about 32% of $TC_{CNS}$. In addition, by adding $TOC_{WB}$ plus $IC_{CARB}$, the quantity of carbon estimated was slightly below of $TC_{CNS}$ content. According to resulted data, it was possible to consider that the organic carbon expressed by both $OC_{CNS}$ and $TOC_{WB}$ accounted for 47% and 38% of LOI fraction respectively, showing moreover a large difference in terms of $OC_{CNS}$ or $TOC_{WB}$/LOI

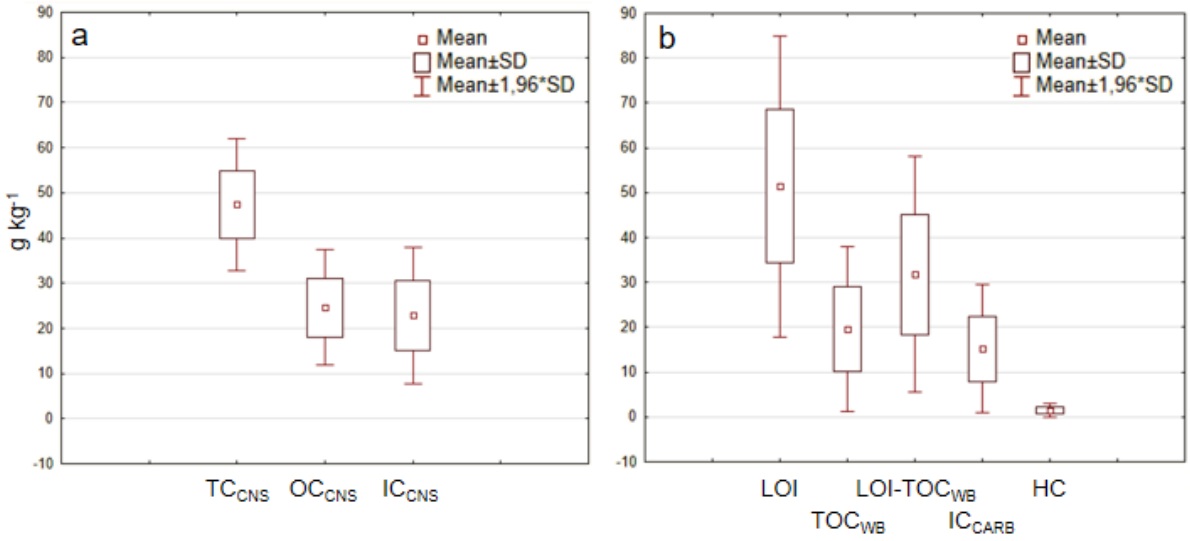


Figure 3.1 Box plots for the partitioning of a) total carbon measured with CNS ($TC_{CNS}$), which comprises the fractions HCl treated ($OC_{CNS}$) and the difference ($TC_{CNS}$ - $OC_{CNS}$) as inorganic carbon ($IC_{CNS}$); and b) LOI, total organic carbon by wet oxidation ($TOC_{WB}$), hot water extractable organic carbon (HC) and inorganic carbon with HCl ($IC_{CARB}$).

ratio values according to the sampled sites. Both $OC_{CNS}$ and $TOC_{WB}$ showed an important correlation with LOI, which corroborates our findings (Figure 3.2). The difference in the correlation coefficients for both organic carbon fractions could be due to differences in clay mineralogy causing differences in structurally bound water and might be related to losses other than SOM fractions. The organic oxidized fraction ($TOC_{WB}$) was 38% of LOI while the SOC quantified by high-temperature dry combustion methods ($OC_{CNS}$) was about the 47% indicating that the conventional LOI-to-SOC conversion factor of 0.58 (Pribyl, 2010) was not applicable in our soils. These results seemed to indicate that a fair difference (about 20%) between the methods of high temperature dry combustion ($OC_{CNS}$) and dichromate oxidation ($TOC_{WB}$) could be related to the presence of resistant organic materials as well as water lost from the mineral phases. The smallest organic carbon fraction, as hot water-extractable carbon (HC), accounted for about 6% of total $OC_{CNS}$ and 7.5% of $TOC_{WB}$.

Moreover, the inorganic carbon expressed as carbonate ($IC_{CARB}$), could be considered the carbon fraction that highly varied among sites according the dynamic of the lixiviation of the carbonate.

Scatter plots revealed the strong relationships between measured data for different SOC fractions (Figure 3.2). Outliers were not deleted because the study was based on 56 samples with a large variability in the topographic position and parent material.

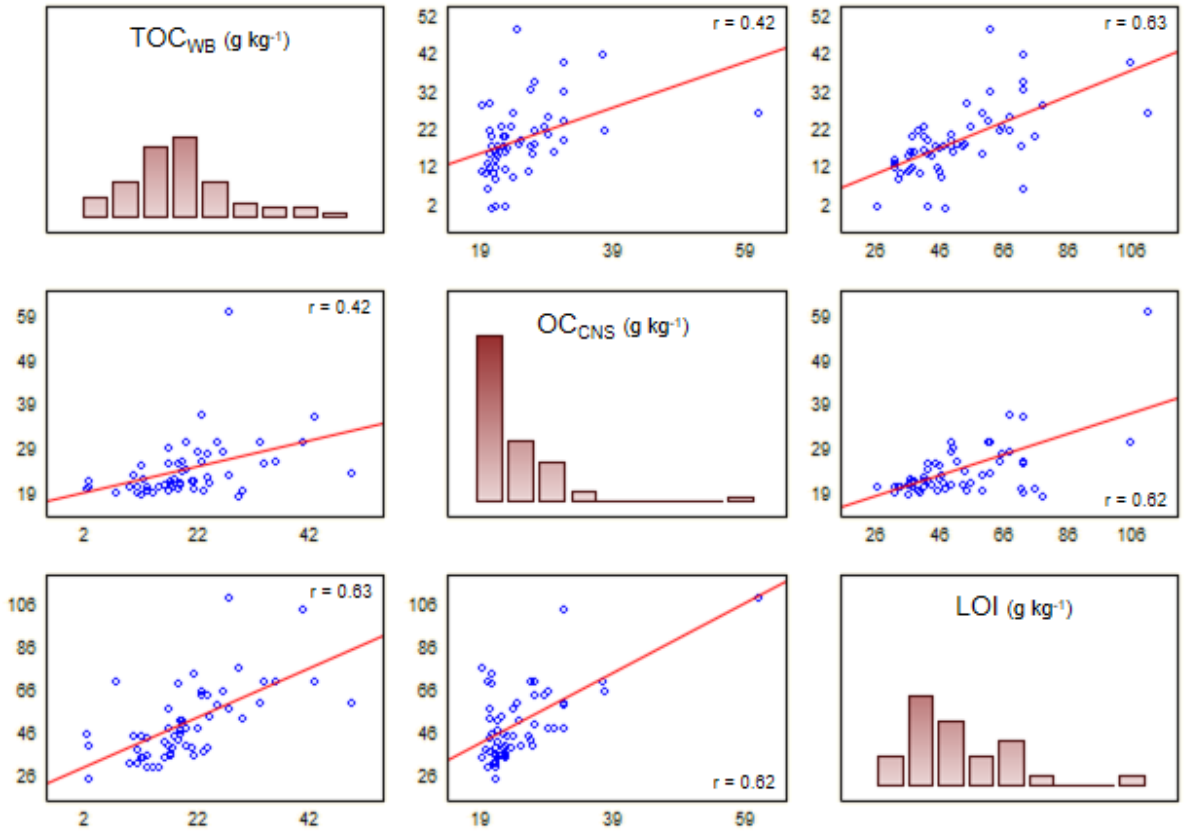


Figure 3.2 Pearson correlations (*p*-value <0.05) and frequency distribution between measured soil organic carbon ($OC_{CNS}$ and $TOC_{WB}$) and loss-on-ignition (LOI).

### 3.1.2 Soil properties affecting the organic and inorganic carbon fractions

Both the organic carbon $OC_{CNS}$ and $TOC_{WB}$ fractions, although obtained with two completely different methods, were considered the less recalcitrant soil C pool and were composed of complex organic molecules with high degree of decomposability by microorganisms. While the LOI-$TOC_{WB}$ and $IC_{CNS}$ fractions could contain a lower part of organic molecules associated with clay and silt or stabilized in aggregates together with inert carbon (i.e. $CaCO_3$) with low degree of decomposability. We considered this fraction as the high-term accumulation of soil organic matter and the most important carbon fraction from a sequestration perspective (Ontl and Schulte, 2012).

Although conducted for several consecutive years (2016, 2017, 2018), the analysis of the soil organic carbon content for the studied area did not show evident differences in terms of accumulation of organic matter (SOM) in the studied area. This results suggested that the interval time for observing the SOM quantitative variations in the field is too limited, while we observed that the variations between the organic and inorganic carbon fractions was very significant. Therefore, it was decided to try to highlight the spatial variations of the different

Table 3.1
Basic statistics of the analyzed soil variables (n = 56).

| | Min | Max | Mean | SD | Variance | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|
| $TOC_{WB}$ (g $kg^{-1}$) | 1.93 | 49.12 | 19.59 | 9.34 | 87.30 | 0.83 | 1.46 |
| $TC_{CNS}$ (g $kg^{-1}$) | 32.18 | 70.05 | 47.44 | 7.50 | 56.22 | 0.46 | 1.24 |
| $OC_{CNS}$ (g $kg^{-1}$) | 18.96 | 60.48 | 24.65 | 6.51 | 42.32 | 3.40 | 16.26 |
| HC (g $kg^{-1}$) | 0.52 | 4.17 | 1.47 | 0.79 | 0.62 | 1.51 | 2.26 |
| LOI (%) | 2.59 | 11.04 | 5.14 | 1.72 | 2.93 | 1.33 | 2.39 |
| $IC_{CNS}$ (g $kg^{-1}$) | 0 | 38.80 | 22.86 | 7.67 | 58.78 | -1.19 | 2.45 |
| TN (g $kg^{-1}$) | 1.63 | 6.99 | 3.52 | 1.05 | 1.11 | 0.98 | 1.36 |
| C/N | 3.68 | 15.50 | 7.39 | 2.11 | 4.44 | 1.22 | 3.10 |
| $CaCO_3$ (%) | 0 | 22.33 | 13.83 | 6.56 | 43.01 | -1.13 | -0.09 |
| $CaCO_3$ act (%) | 0 | 15.35 | 5.65 | 3.98 | 15.87 | 0.65 | 0.12 |
| pH | 6.11 | 7.21 | 6.93 | 0.22 | 0.05 | -1.85 | 3.71 |
| EC (µS $cm^{-1}$) | 102 | 147.30 | 123.21 | 7.98 | 63.70 | 0.32 | 1.09 |
| CEC (cmol $kg^{-1}$) | 9.36 | 25.96 | 19.50 | 2.76 | 7.60 | -0.23 | 3 |
| $Na_{exg}$ (cmol $kg^{-1}$) | 0.11 | 1.40 | 0.61 | 0.43 | 0.18 | 0.38 | -1.32 |
| $Ca_{exg}$ (cmol $kg^{-1}$) | 17.84 | 42.02 | 29.40 | 6 | 36.06 | -0.10 | -0.85 |
| $Mg_{exg}$ (cmol $kg^{-1}$) | 0.58 | 4.31 | 2.50 | 0.65 | 0.42 | 0.05 | 0.84 |
| $K_{exg}$ (cmol $kg^{-1}$) | 0.19 | 1.14 | 0.49 | 0.16 | 0.03 | 1.39 | 4.46 |
| $P_{ass}$ (mg $kg^{-1}$) | 0 | 26.09 | 2.99 | 4.16 | 17.32 | 3.37 | 16.69 |
| $Fe_{DTPA}$ (mg $kg^{-1}$) | 8.95 | 67.70 | 21.19 | 7.62 | 58.09 | 4.12 | 25.38 |
| $Mn_{DTPA}$ (mg $kg^{-1}$) | 10.11 | 51.17 | 22.41 | 10.19 | 103.84 | 1.52 | 1.79 |
| $Cu_{DTPA}$ (mg $kg^{-1}$) | 2.59 | 23.79 | 11.61 | 5.81 | 33.75 | 0.23 | -0.96 |
| $Zn_{DTPA}$ (mg $kg^{-1}$) | 0.55 | 13.43 | 5.66 | 3.63 | 13.14 | 0.32 | -0.94 |
| WC (%) | 12.62 | 47.22 | 29.09 | 5.53 | 30.59 | 0.45 | 2.26 |
| WHC (%) | 27.55 | 58.84 | 40.86 | 7.47 | 55.84 | 0.34 | -0.52 |
| Sand (%) | 25.09 | 59.93 | 42.70 | 5.07 | 25.67 | 0.28 | 5.29 |
| Silt (%) | 17.04 | 42.21 | 28.34 | 4.45 | 19.76 | 0.15 | 1.80 |
| Clay (%) | 18.13 | 35.78 | 28.96 | 3.75 | 14.03 | -0.29 | -0.06 |
| BD (g $cm^{-3}$) | 0.84 | 1.47 | 1.19 | 0.12 | 0.01 | -0.83 | 2.49 |
| Porosity (%) | 44.59 | 68.56 | 55.27 | 4.41 | 19.44 | 0.83 | 2.49 |
| Stock C (kg $m^{-2}$) | 0.46 | 9.56 | 4.45 | 1.80 | 3.26 | 0.19 | 0.73 |
| Resp (mg $CO_2$ $g^{-1}$ $d^{-1}$) | 0.03 | 0.45 | 0.24 | 0.09 | 0.01 | 0.13 | -0.13 |
| Cmic (mg $g^{-1}$) | 1.36 | 1.86 | 1.56 | 0.07 | 0.01 | 1.44 | 5.49 |
| Ksat (cm $d^{-1}$) | 185.76 | 1817.64 | 753.20 | 330.67 | 109344.70 | 1.22 | 1.83 |

organic carbon fractions, in relation not solely to the other properties of the soil (physical, chemical and biological), but also to topographical and pedogenetic factors (spatial variations).

Despite the high spatial variability of the organic carbon content ($OC_{CNS}$ and $TOC_{WB}$), which will be shown later, its values appeared to be high, as best explained by the calculation of the mean + or – 1.96 SD (Figure 3.1). In comparison with other tilled soil, as reported in different studies (Ogle et al., 2005; Haddaway et al., 2016), obtained high values of SOM could be attributed to the presence and management of grass cover. Furthermore, as indicated in Massaccesi et al. (2018), herbaceous species are known to enrich soil with labile organic substances through rhizodeposition processes and activity of the root-associated microorganisms.

The value of HC, which constituted the soluble organic carbon fraction with highest activity and lability, represented on average 7.5% of total organic carbon ($TOC_{WB}$), and it was positively correlated to both $TOC_{WB}$ and $OC_{CNS}$ values, as it appears from the Table 3.2, which shows Pearson correlation matrix between soil properties, with highlighted corresponding *p*-value. The HC fraction represented the soluble carbon in water, and it was considered the readily bioavailable SOM (Marschner and Kalbitz, 2003). HC was also the most mobile fraction of SOM and can therefore reach almost all soil macro and micro aggregates by diffusion and convection. In agricultural soils, dissolved organic carbon (DOC) in field-moist soils was very low with content close to 0.4% of SOC; in forest soils, it is higher with value close to 2% (Haynes, 2005). DOM consists of SOM ranging from small-defined molecules (hexose-to-pentose carbohydrates) indicated predominantly microbial origin and readily mineralizable (Marschner and Kalbitz, 2003).

The parameters that showed a great correlation with organic carbon fractions ($OC_{CNS}$, $TOC_{WB}$ and LOI) were C inorganic fractions ($IC_{CNS}$ and $IC_{CARB}$) whit a negative correlation, total nitrogen, exchangeable bases (Ca, Mg and K), water content, porosity, Ksat, soil respiration and microbial activity ($qCO_2$).

As shown in Table 3.1, total $CaCO_3$ content ranged from 0 to 22.33%. In this case, Skewness index (value of -1.13) described a distribution that was very different from normality; moreover, the frequency distribution highlighted a bimodal trend, which suggested the presence of two distinct population of samples according to this parameter. Indeed, samples located in the southern plot of the study area (refer to Figure 2.4) were completely decarbonated: because of they were positioned on two different slopes, with different aspect and slopes, it suggested that the soils located on them has been originated from different parent material, which could explain the high variation in total carbonate content. Indeed, as resulted from geological map of Molise region (Vezzani et al., 2004) and as it has been said in section 2.1, the geological substrate of the territory of Castel del Giudice is characterized by the dominance of the silico-

Table 3.2
Pearson correlation coefficients between measured soil properties (correlations with a *p*-value <0.05 are in bold).

| | $TOC_{WB}$ | $TC_{CNS}$ | $OC_{CNS}$ | HC | LOI | LOI-$TOC_{WB}$ | $IC_{CARB}$ | $TOC_{WB}$+ $IC_{CARB}$ | $IC_{CNS}$ | TN | C/N | $CaCO_3$ | $CaCO_3$ act | pH | EC | CEC | Ca | Mg | K | P | Fe | Mn | Cu | Zn | WC | WHC | Sand | Silt | Clay | BD | Porosity | stock C | Resp | $C_{MIC}$ | $qCO_2$ | CEM | Ksat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $TOC_{WB}$ | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $TC_{CNS}$ | 0.244 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $OC_{CNS}$ | **0.421** | **0.371** | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| HC | **0.355** | 0.018 | **0.302** | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| LOI | **0.626** | 0.251 | **0.622** | **0.598** | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| LOI-$TOC_{WB}$ | 0.103 | 0.150 | **0.501** | **0.516** | **0.840** | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $IC_{CARB}$ | **-0.342** | **0.448** | **-0.290** | **-0.513** | **-0.645** | **-0.585** | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $TOC_{WB}$+$IC_{CARB}$ | -0.102 | **0.485** | **0.423** | **0.292** | **0.603** | **0.840** | -0.051 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| $IC_{CNS}$ | -0.111 | **0.675** | **-0.434** | -0.215 | -0.250 | -0.242 | **0.665** | 0.148 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TN | **0.329** | **0.519** | **0.517** | **0.279** | **0.647** | **0.597** | -0.114 | **0.659** | 0.101 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| C/N | **0.806** | -0.086 | 0.153 | 0.191 | 0.256 | -0.234 | **-0.304** | **-0.492** | -0.222 | -0.247 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | |
| $CaCO_3$ | **-0.333** | **0.445** | **-0.281** | **-0.516** | **-0.645** | **-0.591** | **0.999** | -0.059 | **0.655** | -0.113 | **-0.297** | 1 | | | | | | | | | | | | | | | | | | | | | | | | |
| $CaCO_3$ act | **-0.351** | **0.290** | -0.234 | **-0.459** | **-0.503** | **-0.398** | **0.785** | 0.036 | **0.468** | -0.038 | **-0.346** | **0.788** | 1 | | | | | | | | | | | | | | | | | | | | | | | |
| pH | **-0.272** | **0.278** | -0.231 | **-0.467** | **-0.668** | **-0.663** | **0.830** | -0.262 | **0.452** | -0.235 | -0.171 | **0.834** | **0.650** | 1 | | | | | | | | | | | | | | | | | | | | | | |
| EC | **0.305** | **0.578** | 0.185 | 0.017 | 0.194 | 0.035 | **0.352** | **0.279** | **0.410** | **0.478** | -0.010 | **0.353** | **0.361** | **0.319** | 1 | | | | | | | | | | | | | | | | | | | | | |
| CEC | 0.110 | -0.208 | 0.023 | 0.135 | 0.121 | 0.077 | **-0.272** | -0.087 | -0.221 | 0.001 | 0.120 | **-0.270** | -0.248 | -0.218 | **-0.265** | 1 | | | | | | | | | | | | | | | | | | | | |
| Ca | **0.557** | 0.134 | **0.382** | 0.254 | **0.596** | **0.373** | **-0.366** | 0.215 | -0.187 | **0.361** | **0.352** | **-0.361** | **-0.365** | **-0.357** | **0.331** | -0.023 | 1 | | | | | | | | | | | | | | | | | | | |
| Mg | **0.297** | -0.190 | 0.256 | **0.455** | **0.606** | **0.567** | **-0.657** | 0.258 | **-0.393** | 0.245 | 0.190 | **-0.659** | **-0.668** | **-0.653** | -0.112 | 0.038 | **0.654** | 1 | | | | | | | | | | | | | | | | | | |
| K | **0.303** | 0.162 | 0.149 | 0.062 | **0.265** | 0.127 | -0.002 | 0.156 | 0.031 | **0.345** | 0.129 | -0.001 | 0.001 | -0.130 | **0.299** | 0.075 | **0.677** | **0.305** | 1 | | | | | | | | | | | | | | | | | |
| P | -0.127 | 0.164 | -0.015 | -0.258 | -0.155 | -0.109 | **0.359** | 0.106 | 0.168 | 0.154 | -0.219 | **0.359** | 0.257 | **0.264** | **0.368** | -0.013 | 0.240 | -0.029 | **0.551** | 1 | | | | | | | | | | | | | | | | |
| Fe | 0.128 | -0.176 | -0.005 | **0.310** | **0.312** | **0.309** | **-0.353** | 0.145 | -0.166 | 0.169 | 0.009 | **-0.353** | **-0.344** | **-0.570** | -0.160 | 0.170 | 0.062 | **0.441** | 0.000 | -0.150 | 1 | | | | | | | | | | | | | | | |
| Mn | -0.012 | -0.071 | 0.136 | **0.275** | **0.296** | **0.386** | **-0.299** | **0.276** | -0.176 | 0.142 | -0.092 | **-0.301** | **-0.317** | **-0.475** | -0.250 | 0.215 | 0.089 | **0.387** | 0.247 | 0.071 | **0.600** | 1 | | | | | | | | | | | | | | |
| Cu | -0.102 | -0.073 | 0.026 | -0.236 | -0.243 | -0.239 | 0.254 | -0.125 | -0.101 | -0.102 | -0.078 | **0.266** | 0.228 | 0.218 | 0.055 | -0.068 | 0.025 | -0.111 | 0.197 | 0.159 | -0.021 | 0.198 | 1 | | | | | | | | | | | | | |
| Zn | -0.086 | -0.082 | 0.057 | -0.253 | -0.201 | -0.196 | 0.195 | -0.111 | -0.135 | -0.050 | -0.085 | 0.208 | 0.218 | 0.183 | 0.046 | -0.058 | 0.050 | -0.112 | 0.218 | 0.158 | -0.099 | 0.148 | **0.972** | 1 | | | | | | | | | | | | |
| WC | **0.278** | 0.202 | 0.238 | -0.063 | **0.282** | 0.167 | -0.090 | 0.145 | -0.002 | 0.230 | 0.160 | -0.084 | 0.005 | 0.019 | 0.121 | -0.210 | 0.233 | -0.030 | 0.110 | -0.118 | **-0.308** | -0.012 | 0.223 | **0.271** | 1 | | | | | | | | | | | |
| WHC | 0.211 | -0.063 | 0.146 | **0.384** | **0.377** | **0.334** | **-0.324** | 0.195 | -0.178 | 0.052 | 0.178 | **-0.324** | -0.136 | **-0.283** | -0.099 | -0.068 | 0.162 | 0.216 | -0.041 | **-0.306** | 0.206 | 0.164 | -0.049 | -0.001 | **0.265** | 1 | | | | | | | | | | |
| Sand | -0.116 | -0.089 | -0.050 | 0.101 | -0.144 | -0.104 | -0.072 | -0.176 | -0.043 | -0.246 | 0.002 | -0.074 | -0.206 | 0.027 | -0.240 | -0.065 | -0.221 | -0.095 | -0.177 | -0.050 | 0.003 | -0.045 | -0.035 | -0.055 | -0.108 | -0.188 | 1 | | | | | | | | | |
| Silt | 0.191 | **0.271** | 0.009 | -0.052 | 0.025 | -0.100 | 0.255 | 0.047 | 0.251 | 0.101 | 0.155 | 0.258 | 0.217 | 0.210 | 0.231 | 0.048 | -0.032 | -0.105 | -0.101 | -0.088 | -0.099 | -0.064 | -0.086 | -0.114 | 0.213 | 0.174 | **-0.697** | 1 | | | | | | | | |
| Clay | -0.070 | -0.201 | 0.057 | -0.074 | 0.165 | 0.260 | -0.206 | 0.182 | -0.239 | 0.213 | -0.187 | -0.206 | 0.021 | **-0.285** | 0.051 | 0.031 | **0.337** | 0.254 | **0.358** | 0.172 | 0.114 | 0.137 | 0.149 | 0.209 | -0.108 | 0.047 | **-0.525** | -0.244 | 1 | | | | | | | |
| BD | **-0.297** | -0.142 | -0.143 | -0.127 | -0.123 | 0.049 | -0.013 | 0.052 | -0.016 | -0.119 | -0.194 | -0.028 | 0.050 | -0.027 | -0.199 | 0.156 | -0.216 | -0.086 | -0.058 | -0.018 | **-0.278** | -0.119 | -0.043 | 0.066 | -0.124 | -0.077 | 0.074 | -0.107 | 0.026 | 1 | | | | | | |
| Porosity | **0.297** | 0.142 | 0.143 | 0.127 | 0.123 | -0.049 | 0.013 | -0.052 | 0.016 | 0.119 | 0.194 | 0.028 | -0.050 | 0.027 | 0.199 | -0.156 | 0.216 | 0.086 | 0.058 | 0.018 | **0.278** | 0.119 | 0.043 | -0.066 | 0.124 | 0.077 | -0.074 | 0.107 | -0.026 | **-0.999** | 1 | | | | | |
| stock C | **0.875** | **0.329** | **0.408** | **0.384** | **0.656** | 0.228 | **-0.304** | 0.078 | -0.013 | **0.366** | **0.649** | **-0.306** | **-0.332** | **-0.306** | **0.361** | 0.088 | **0.511** | **0.285** | 0.203 | -0.093 | 0.125 | -0.022 | -0.159 | -0.134 | 0.244 | 0.226 | -0.064 | 0.158 | -0.101 | -0.020 | 0.020 | 1 | | | | |
| Resp | **0.512** | -0.027 | **0.399** | **0.381** | **0.694** | **0.528** | **-0.622** | 0.234 | **-0.352** | **0.407** | **0.284** | **-0.617** | **-0.423** | **-0.630** | 0.101 | -0.045 | **0.643** | **0.531** | **0.415** | -0.191 | 0.220 | **0.268** | 0.062 | 0.131 | **0.413** | **0.349** | -0.075 | -0.085 | 0.202 | -0.006 | 0.006 | **0.478** | 1 | | | |
| $C_{MIC}$ | -0.132 | 0.054 | -0.027 | -0.113 | -0.077 | -0.006 | 0.135 | 0.083 | 0.077 | 0.056 | -0.215 | 0.137 | **0.424** | 0.034 | **0.273** | **-0.311** | 0.087 | **-0.294** | 0.147 | 0.077 | -0.071 | -0.039 | **0.340** | **0.345** | 0.151 | 0.041 | -0.125 | -0.177 | **0.380** | -0.103 | 0.103 | -0.063 | 0.226 | 1 | | |
| $qCO_2$ | **0.551** | -0.027 | **0.412** | **0.409** | **0.722** | **0.537** | **-0.652** | 0.225 | **-0.364** | **0.411** | **0.329** | **-0.647** | **-0.489** | **-0.645** | 0.068 | -0.010 | **0.649** | **0.584** | **0.412** | -0.208 | 0.230 | **0.278** | 0.009 | 0.080 | **0.407** | **0.356** | -0.063 | -0.059 | 0.154 | 0.002 | -0.002 | **0.500** | **0.989** | 0.086 | 1 | |
| CEM | **-0.569** | -0.085 | -0.162 | -0.093 | -0.083 | **0.290** | 0.030 | **0.377** | 0.052 | 0.128 | **-0.618** | 0.028 | 0.238 | -0.063 | -0.099 | -0.180 | -0.072 | 0.045 | 0.014 | -0.030 | -0.009 | 0.093 | -0.048 | -0.008 | -0.067 | 0.000 | -0.055 | -0.149 | 0.251 | 0.082 | -0.082 | **-0.616** | 0.047 | 0.167 | 0.024 | 1 |
| Ksat | **-0.375** | -0.050 | -0.149 | -0.074 | **-0.314** | -0.139 | 0.179 | -0.051 | 0.071 | -0.190 | **-0.306** | 0.186 | 0.070 | 0.186 | -0.099 | -0.230 | -0.167 | -0.097 | -0.108 | 0.103 | 0.143 | 0.063 | 0.047 | -0.060 | -0.164 | -0.197 | **0.386** | **-0.360** | -0.095 | **-0.656** | **0.656** | **-0.561** | **-0.314** | 0.114 | **-0.335** | **0.298** | 1 |

clastic clay formation, indicated in the literature as the Flysch of Agnone. However, the whole unit is characterized by frequent intercalations of a calcarenite-calciruditic nature. Therefore, given the chaotic tectonics of the territories, the two different formations can be found side by side even within a few meters.

Anyway, with a mean value of 13.83% carbonate content, the studied soils could be consider as moderately calcareous soils. Active calcium carbonate, significantly and positively correlated to total carbonate (Table 3.2), represented the fraction that reacts more readily with the other components of the soil; it influenced the availability of phosphorus and iron forming with them high insoluble compounds that cannot be assimilated by the plant. Here, the presence of high percentages of active calcium carbonate (maximum of 15.35%) indicated an unfavorable situation in terms of nutrient acquisition by apple trees, for which M9 rootstocks a threshold value of 10% was indicated, beyond which find damage to tree crops. Indeed, as resulting from correlation matrix (Table 3.2), active carbonate was positively correlated to soil pH, for which a mean value of 6.93 and a low standard deviation (0.22) indicated a completely neutral reaction.

Conversely, active carbonate content was negatively correlated to exchangeable Ca and Mg and assimilable Fe and Mn, but did not show evident correlation to value of assimilable phosphorus. This last parameter ranged from 0 to 26.09 mg $kg^{-1}$, with a mean of 2.99 mg $kg^{-1}$, values that indicated a very low assimilable phosphorus content, as could be expected given the high concentration of active calcium carbonate.

Measures about exchangeable Ca, Mg and K contents, whose mean values were 29.4, 2.5 and 0.49 cmol $kg^{-1}$ respectively, indicated a general good supply of these elements (MiPAF, 1999), but, the Mg/K ratio (mean value is 5.43) may consist in an antagonistic effect of magnesium on potassium absorption (Laekemariam et al., 2018). Instead, as regards to microelements of fertility, results of analysis on available Fe, Mn, Cu and Zn contents, showed a rich supply of these elements (Sequi, 2010), the average values of which amounted to 21.19, 22.41, 11.61 and 5.66 g $kg^{-1}$ respectively.

Soil BD ranged between 0.88 and 1.47 g $cm^{-3}$ (mean 1.19 g $cm^{-3}$) and was significantly lower in soils with higher total carbon ($TOC_{WB}$). There was no statistically significant difference in BD between the topsoils analyzed. Lastly, the hydraulic properties of the studied soils could be evaluated according to the value of the corresponding Ksat index, which indicated the rate of infiltration of water into the soil, and therefore was related to the main soil physical properties. First of all, according to USDA system classification (SSDS, 2017), Ksat, whose mean value accounted for 753.20 cm $d^{-1}$ and ranged from 185.76 to 1817.64 cm $d^{-1}$, could be consider high

or very high. This was probably due to the presence of grass cover roots and good amount of soil organic matter in the top soil layer. Again, as an expected implication, Ksat was positively correlated to sand concentration and porosity, and negatively to contents of smaller soil particles, in particular silt, and values of bulk density.

### 3.1.3 Soil properties affecting the labile organic carbon fraction and biological properties

As an expected result, values of HC for the studied soils positively correlated with organic carbon fractions and several biological properties. In particular, the increasing of labile pool of organic carbon resulted in an increase of microbial activity, expressed by levels of $qCO_2$ index, which indicated the part of $CO_2$ actually produced by the microorganisms during the incubation period to which the soils have been subjected. The other biological index (CEM), which explained the rate of organic carbon mineralization, according to Table 3.2, was negatively correlated to C/N ratio and organic carbon stock content. This kind of situation was due to the quality of SOM in the soil, according to C/N ratio: as this value increased, the degradability of the organic matter decreased, and consequently its mineralization rate.

Total nitrogen (TN) content, obtained by CNS analyzer, ranged from 1.63 to 6.99 g $kg^{-1}$, with a mean value of 3.52 g $kg^{-1}$, which suggested an excellent supply of soil nitrogen, but drastically decreased the C/N value, which was at an average of 7.39. With this value of C/N, oxidation reactions prevailed, resulting in a decrease in the organic matter content and in release of assimilable nitrogen. Mineralization rate of SOM was also influenced by clay and total carbonate contents. As resulted by soil texture analysis and shown in Figure 3.3 for clarity, the top soil samples had three types of soil textures: sandy clay loam, clay loam and loam. From Table 3.2, clay content was positively correlated to $C_{MIC}$ property, and this suggested increased microbial activity as clay content increased, further explained by a consecutive decrease in the rate of SOM mineralization (Schillaci et al., 2017). On the other hand, both mineralization (Resp) and microbial activity ($qCO_2$) significantly and negatively correlated to total carbonate content, this was because the greater the amount of calcium carbonate, the greater the inertia of the soil towards the transformation processes of organic compounds.

### 3.1.4 The organic carbon accumulation and soil properties

The dominant effect in the studied soils was the development of porosity as indicated by the significant reduction in BDs in soil with higher $OC_{CNS}$ and $TOC_{WB}$. The value of carbon accumulated (Stock C kg $m^{-2}$) in the top-soil (20 cm) ranged from 0.46 to 9.56 kg $m^{-2}$ (mean

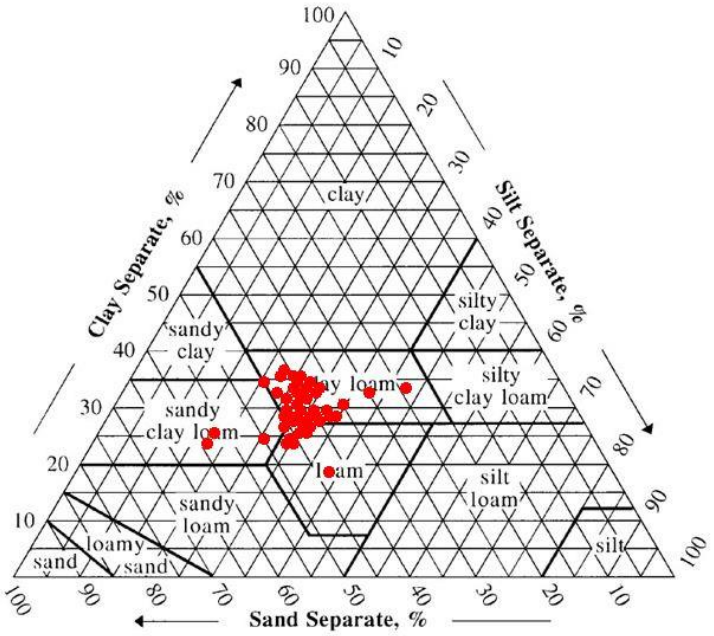


Figure 3.3 Soil texture contents of samples (red dots) according to USDA classification system.

4.45 stock C kg $m^{-2}$) and seemed to increase the formation of macro-pores and macro-aggregates induced by the cementing action with the manure-fertilized treatments.

These results could reflect higher levels of stubble, root and rhizodeposition inputs that could be attributed to the presence and management of grass cover. This condition improved the labile organic carbon (HC) through rhizodeposition processes and activity of the root-associated microorganisms (Kalcsits et al., 2020). This pattern reflected strongly changes in the SOC variability; highest HC plot was correlated with the lowest calcium carbonate and higher LOI. The amount of C sequestered and the rate of C sequestration seemed related to the topographic and exposition conditions. The labile pools could not benefit C sequestration because they were rapidly mineralized by soil organisms (Stockmann et al., 2013).

In this research, SOC concentrations in organic management soil were greatly increased after 10 years of orchard-cropping farming system (no-till approach) compared to initial estimated soil levels (<10 C g $kg^{-1}$). This can be attributed an increase in root exudates from grass cover and from the C contained in the stubble which is returned to the soil each year, and also to the conservation management, which reduces soil disturbance.

Soil C levels in the organic fertilized treatments also increased marginally relative to the untreated soils. There are conflicting reports about the impact of agriculture management practices on SOC sequestration (Lal, 2004; 2008); some reports demonstrated that the change in land use induced a net loss of SOC, however, our results indicated that long-term orchard-cropping farming system increase SOC stocks in the topsoil layer. The C fractions differences

seemed to depend on the initial soil C status, the soil pedology properties under study (i.e. the physical and chemical driving factors), the quantity and quality of residues returned and the nature, quantity and duration of fertilizer application (Li et al., 2018). There were no statistical significant differences between SOC and C stock concentrations vs. texture (clay content), pH and CEC, which were not linked with chemical properties but with crop residues in these soils. This was possibly attributed to a larger proportion of recalcitrant organic compounds accumulated in the topsoils as lignin and lignin-like products, which were the major components of the resistant C pool in the soil (Schmidt et al., 2011).

Microbial carbon ($C_{MIC}$) and HC were produced from decomposition of soil organic matter mainly driven by soil microbes (Lima et al., 2009). They were well correlated with LOI, soil respiration (Resp) and microbial activity ($qCO_2$), highly correlated each other and with the organic carbon content and stock, expressed as $OC_{CNS}$/$TOC_{WB}$ and stock C, respectively. Although $C_{MIC}$ and HC accounted for only a small proportion of organic carbon stock (generally 0.80–12.00% for $C_{MIC}$) in agricultural soils (Ding et al., 2012), these measures could be considered good indicators of the soil's potential C accumulation, a key ecosystem service (Haynes, 2005).

### 3.1.5 Results of principal component analysis

Principal component analysis was used to reveal the hidden structure within dataset. It provides a visual representation of the relationships between samples and variables and provides new insights into how measured variables cause some samples to be similar to, or how they differ from each other. Using standardized physico-chemical and biological soil properties as continuous variables in PCA, based on Figure 3.4, constructed first and second components (PC1 and PC2) were able to describe 38% of total variance of dataset, 25% from PC1 and 13% from PC2. In particular, score plot (Figure 3.4.a) revealed weakly clusters in the observed dataset, according soil properties. At the same time, Hotelling's $T^2$ ellipse, used as a valid tool to detect outliers, didn't show anomalous observations, assuming then that all the soil samples were part of the same population, despite the descriptive statistics evaluated above showed a bimodal trend for the total calcium carbonate content.

Figure 3.4b, shows correlation loadings from PCA, and allows to detect important variables according to PC1 and PC2. Soil variables that are located outside the small ellipse in the graph contributed more than 50% to the construction of the new component system in the PC analysis, until a percentage of 100% at the major ellipse. In other words, it is a graphical method that, in

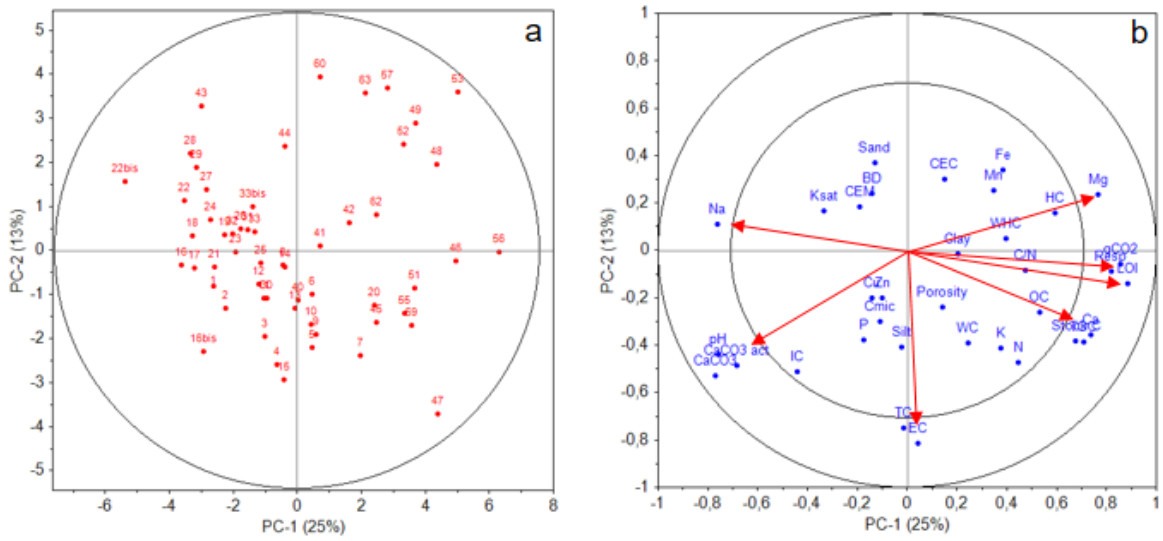


Figure 3.4 Plots resulting from principal component analysis: a) score plot of 56 topsoil samples (red dots) according to PC1 and PC2, with highlighted Hotelling's $T^2$ ellipse for outliers detection; and b) correlation loadings plot of soil physico-chemical and biological properties (X-loadings). The two ellipses indicate how much variance is taken into account (100 and 50%, from outside to inside respectively).

addition to providing information on the importance of the variables, facilitates and corroborate of the trend obtained with the correlation matrix previously shown in Table 3.2. Clarified by the presence of the red arrows in the figure, total carbon ($TC_{CNS}$) and electron conductivity (EC) properties were explained only by PC2, all other important variables from PC1.

In particular, organic matter content expressed by LOI, soil respiration (Resp) and microbial activity ($qCO_2$), highly correlated each other, strongly and negatively correlated to exchangeable Na content, as well as exchangeable Mg and Ca, and organic carbon content and stock, expressed as $TOC_{WB}$ and stock C respectively. The others soil parameters that affected negatively on Mg, Ca, organic carbon contents and microbial activity are total and active calcium carbonate, both positively correlated to pH values. Parameters forming angles of 90° from axes origin are uncorrelated.

Although they were not important in the construction of the new PC1 - PC2 projection, the physical-hydraulic properties (sand content, buld density and ksat) closely correlated with each other, demonstrating that as the sandy fraction increased, which had a close link with the parent material, increased the value of BD and ksat, decreasing the soil's ability to hold water. Indeed, both the water content and the WHC were negatively correlated with the physical properties and positively with the SOC content ($OC_{CNS}$, $TOC_{WB}$ and LOI), and therefore the microbial activity (Resp and $qCO_2$) of the soils.

### 3.2 Spatial prediction of soil carbon pools

To evaluate spatial variability of soil carbon pools and analyze soil properties two different semivariogram models, spherical and exponential, were applied. In particular, spherical semivariogram model shows a progressive decrease of spatial autocorrelation (equivalently, an increase of semivariance) until some distance, beyond which autocorrelation is zero. Instead, with exponential approach, spatial autocorrelation decreases exponentially with increasing distance. Then, the autocorrelation disappears completely only at an infinite distance. Table 3.3 shows the results of application of two geostatistical models, whose performances were evaluated by means of Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) and root mean squared error (RMSE). In particular, to obtain a better information on the spatial variability of soil C pools, prediction maps of the other soil attributes were realized by geostatistical approach. The most relevant soil properties were chosen based on significant Pearson correlations ($p$-value <0.05), shown in the previous paragraph (Table 3.2) and on which was discussed above. As to the selection, the following seven variables were analyzed: total calcium carbonate and clay contents, C/N ratio, bulk density and biological properties, expressed as soil respiration, microbial activity and endogenous organic carbon mineralization.

Prior to application of univariate OK method, data of each variable were standardized, having zero mean and unit standard deviation, based on the assumption that geostatistical analysis is more efficient when done on variables that have normal distributions (Webster and Oliver, 2007), as well as done in Ferré et al. (2018) and Zovko et al. (2018).

As shown in Table 3.3, semivariogram models were chosen for each variable according to values of AIC, BIC and RMSE. In general, there were no great differences in terms of

Table 3.3
Performances of application of geostatistical spherical and exponential models in soil attributes spatial predictions, evaluated by AIC, BIC and RMSE values.

| | Spherical model | | | Exponential model | | |
|---|---|---|---|---|---|---|
| | AIC | BIC | RMSE | AIC | BIC | RMSE |
| $TC_{CNS}$ | -8.13 | -7.89 | 0.21 | -8.41 | -8.17 | 0.20 |
| $OC_{CNS}$ | 13.1 | 13.69 | 0.66 | 12.43 | 13.03 | 0.64 |
| $TOC_{WB}$ | 4.94 | 6.84 | 0.36 | 4.85 | 6.01 | 0.36 |
| LOI | 6.16 | 6 | 0.58 | 4.57 | 4.41 | 0.52 |
| HC | 3.4 | 4.51 | 0.34 | -5.37 | -4.18 | 0.23 |
| $IC_{CNS}$ | 5.72 | 5.56 | 0.57 | 4.45 | 4.29 | 0.52 |
| $CaCO_3$ | 7.41 | 8 | 0.48 | 10.5 | 11.09 | 0.57 |

performances between application of spherical and exponential models. This result was probably due to the application of models always on the same size of samples (56 georeferenced points) and their good distribution, so in this case the differences were attributable only to the diversity of models, as indicated by Lee and Ghosh (2009). Anyway, different authors (Aksoy et al., 2012; de Menezes et al., 2016; Pham, 2019) suggested to choose model with lower or more negative values of both AIC and BIC criterions, as well as value of RMSE, which represented the differences between values predicted by the model and the values observed. For all soil C pools attributes (Table 3.3), exponential univariate OK was chosen, despite $CaCO_3$ experimental data showed a better explanation by spherical model (AIC 7.41, BIC 8, RMSE 0.48), compare to exponential (AIC 10.5, BIC 11.09, RMSE 0.57).

Figure 3.5 and 3.6 show semivariogram models of spatially predicted soil properties. As an expected result, all variograms highlighted a nugget effect, although there were differences between the various soil parameters analyzed. The nugget could be attributed to measurement errors or spatial sources of variation at distances smaller than the sampling interval or both. Generally, experimental variograms were well structured with small nugget effect, except for C/N ratio (~ 50% of sill) and coefficient of endogenous mineralization (CEM) (~ 80% of sill) attributes (Figure 3.6c and f, respectively). A nugget that is large relative to the sill, which is

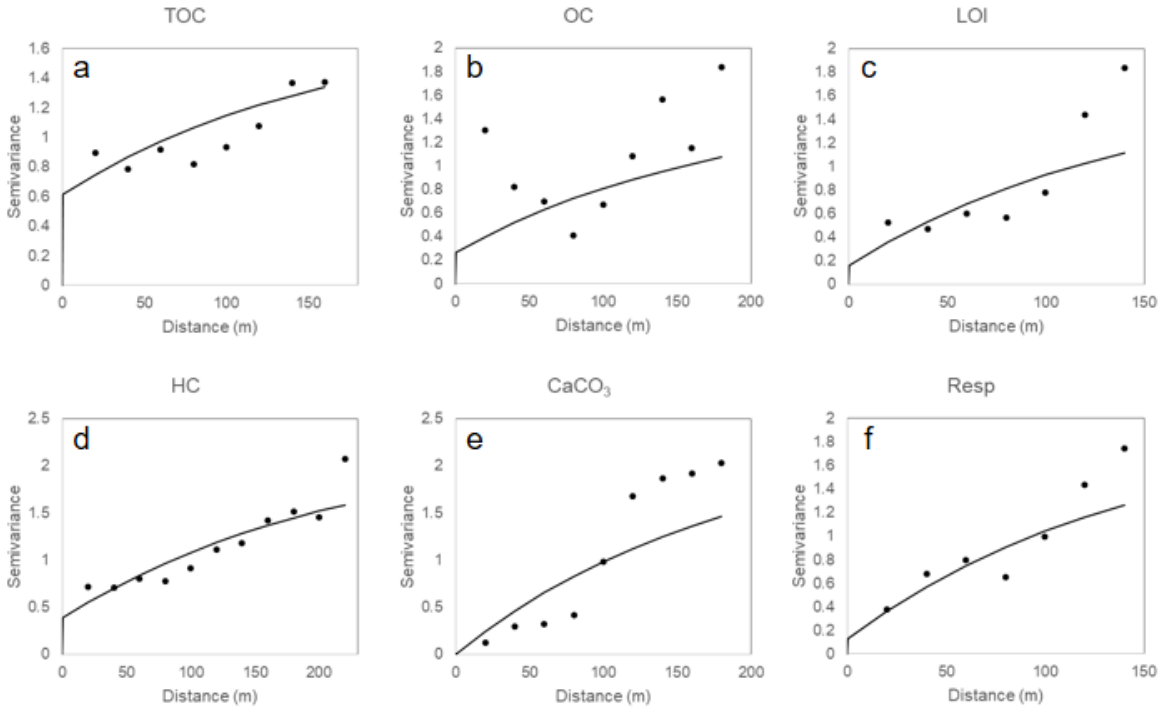


Figure 3.5 Experimental (points) and fitted (solid lines) model semivariograms of measured values of a) total organic carbon ($TOC_{WB}$), b) organic carbon ($OC_{CNS}$), c) loss on ignition (LOI), d) hot water extractable carbon (HC), e) total calcium carbonate content ($CaCO_3$) and f) soil respiration (Resp).

the total variance where the empirical variogram appears to level off, could indicate too much noise and not enough spatial correlation. In particular for CEM (nugget of 1.1), which represented the rate of mineralization of organic carbon in the soil, this probable condition had already been taken over by Pearson correlation matrix (Table 3.2) and validated here with semivariogram model. Therefore, the mineralization rate for the studied soils did not show self-correlation and did not appear to be significantly correlated to the other properties of the soil, as shown additionally in the respective predictive map (Figure 3.8f), in which interpolation of CEM values appeared to be much less regular compare to other soil properties.

As opposite situation, semivariogram of total calcium carbonate ($CaCO_3$) content (Figure 3.5e), which gave a nugget of 0.003, represented the best experimental semivariogram model. As described by de Menezes et al. (2016), variograms were used to describe and measure the spatial structure of spatial data sets.

In case of measured $CaCO_3$ soil property, a nugget to sill ratio (N/S) of 0.002 meant that only 0.2% of the variability consisted of unexplainable or random variation. Additionally, the high variability of $CaCO_3$ contents (see Table 3.1) and the clustering of experimental points in two groups in the semivariogram model (above and below the fitted curve, Figure 3.5e) suggested that soil calcium carbonate spatial content was co-determined by other soil

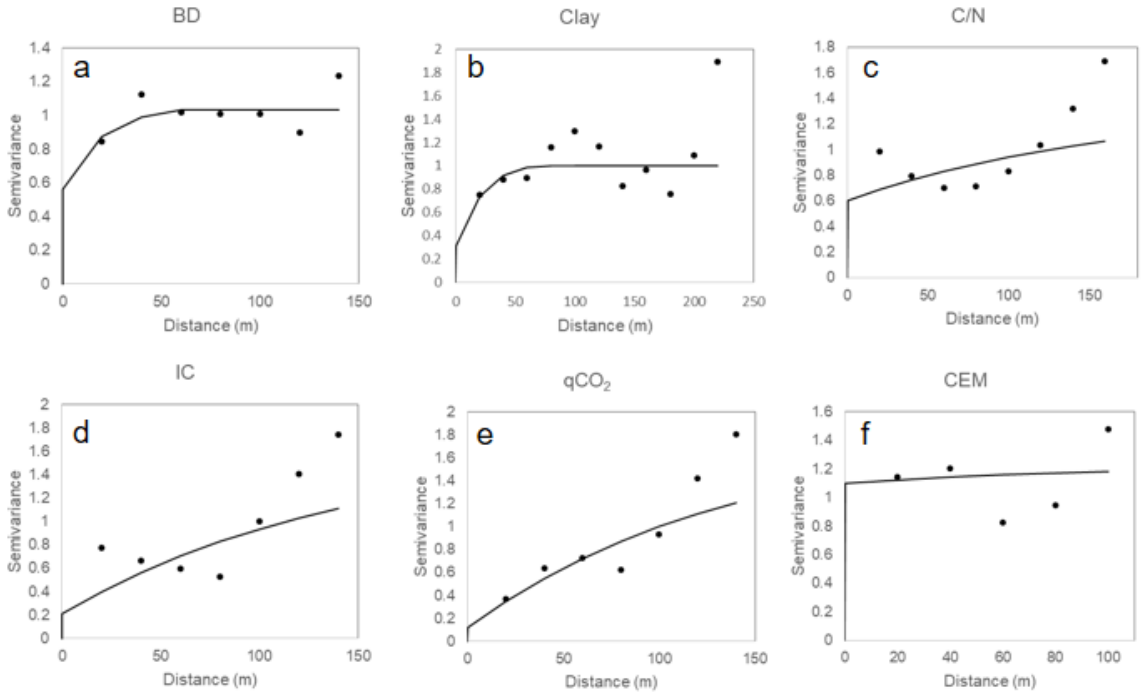


Figure 3.6 Experimental (points) and fitted (solid lines) model semivariograms of measured values of a) bulk density (BD), b) clay content, c) C/N ratio, d) inorganic carbon ($IC_{CNS}$) content, e) microbial activity ($qCO_2$) and f) coefficient of endogenous mineralization (CEM).

properties. $CaCO_3$ content, which behaved then as two distinct populations, was affected with a high probability by structural factors, such as parent material, terrain, mineralogy and pedogenic processes. In turn, the different $CaCO_3$ contents significantly affected the spatial variability of the other soil properties, especially the biological ones.

As resulting from Figure 3.5d, the semivariogram model of hot water extractable carbon (HC) contents, which represented the organic C pool more labile and active in the soil, highlighted the greater range spatial autocorrelation (220 m), which is the distance after which data are no longer correlated, compared to those of the other soil attributes. The presence of this long-range spatial autocorrelation was much probably because, as suggested by Sarkhot et al. (2011), HC values were significantly affected by crop type, quality of C inputs from fine roots etc., all factors that were homogeneous in the studied soils, where permanent coverage was guaranteed by the growth of spontaneous grasses subjected to repeated mowing.

Instead, the semivariogram models of bulk density and clay content (Figure 3.6a and b, respectively) showed a short-range spatial autocorrelations (60 and 80 m, respectively), indicating a strong correlation between them and a strong co-determination for spatial variability from other soil properties, in particular for bulk density values.

Maps of predicted spatial variability of soil organic carbon pools and the other soil attributes strongly correlated, all resulting from application of OK exponential geostatistical model, are shown by Figure 3.7 and 3.8. In particular, Figure 3.7a, b, c and d highlight the distribution of organic C pools. Generally, indicated maps were similar, and the differences among them were in areas with extreme values. Indeed, first study area could be divided in two sub-area: the field to the north, which presented the lower values of organic carbon, and the field to the south, which was characterized by its high organic carbon values. Total organic carbon ($TOC_{WB}$, Figure 3.7a) content, evaluated by Walkley-Black method, showed the major variability in terms of values (from 9.68 to 35.75 g $kg^{-1}$), but varied gradually across the study area and increased from north to south. Instead, organic carbon ($OC_{CNS}$) content (CNS analyzer) and LOI, which accounted with a good approximation to soil organic matter content (Ben-Dor and Banin, 1989), indicated, as the previous variable, an increase of values from north to south. However, differently to $TOC_{WB}$ property, they were able to better highlight the areas with average organic carbon values (23-27 g $kg^{-1}$ for OC and 3.6-5.8 % for LOI; Figure 3.7b and c, respectively), in particular for northern field. This occurrence was also indicated by the semivariogram models of $OC_{CNS}$ and LOI parameters (Figure 3.5b and c, respectively), which showed a short-range autocorrelation, especially for the LOI (range of 140 m). From a topographic point of view, the variability of organic carbon content was in general connected

to soil-landscape relationship (de Menezes et al., 2016). In the studied area, the low difference in altitude (825-852 m) was not sufficient to justify a correlation to the different storage capacity of organic matter of the soils, however this seemed to be linked to their different exposure and slope. As indicated in Figure 3.9a and b, which respectively show the slope and aspect of the studied area, calculated by DEM with an accuracy of 5 m, the field below, in addition to having the greatest slopes (20-25%) compared to the above field, was completely exposed to the north. Therefore, the higher organic matter content detected there was probably due to lower soil temperatures, compared to those of the northern field, which showed a faster organic carbon dynamics, thus proved by the lower accumulations of the $TOC_{WB}$, $OC_{CNS}$ and LOI (Figure 3.7).

From the evaluation of different organic carbon pools in the soil, hot water extractable carbon (HC) was suggested to be a responsive indicator to assess effects of land management and availability of C to microbial population and rate of decomposition (Sarkhot el al., 2011). From Figure 3.7d, HC values had a strong positive spatial correlation to all forms of organic materials ($TOC_{WB}$, $OC_{CNS}$ and LOI). As an expected result, increase of organic carbon content corresponded to an increase of available organic carbon (HC) content. However, the importance of this result was given by % of $HC/TOC_{WB}$ ratio, which showed its greater values of the field to the north than the one to the south, with an opposite trend compared to the values of both TOC and HC. This could mean that the higher percentage of organic carbon more available for the microbial population (HC) compared to the total organic carbon and therefore to the most recalcitrant fractions, further justified a faster carbon dynamics in the northern field. Additionally, as proved by different authors (McLauchlan and Hobbie, 2004; Nawar et al., 2016), C/N ratio, which represented the quality of organic material in the soil, highly contributed to its predisposition to the mineralization. As described by distribution in Figure 3.8c, C/N ratio values, which ranged from 3.7 to 15.5, further testified to a rapid mineralization and therefore a rapid consumption of organic matter in the northern part of the study area, compared to the southern one. Indeed, for those soils, the average C/N values of about 7-8, suggested a greater equilibrium between mineralization and humification phases of the organic matter.

Predictive distributions of organic carbon mineralization, expressed by both soil respiration and microbial activity and shown in Figure 3.7f and 3.4e, respectively, were strongly correlated to the distributions of organic carbon pools. First, this result indicated that all carbon pools ($TOC_{WB}$, $OC_{CNS}$, LOI and HC) were robust predictors of the amount of potentially mineralizable soil organic. However, as suggested by Ahn et al. (2009), variability of soil mineralization rate that seemed to be related in the studied soils only to organic carbon content,

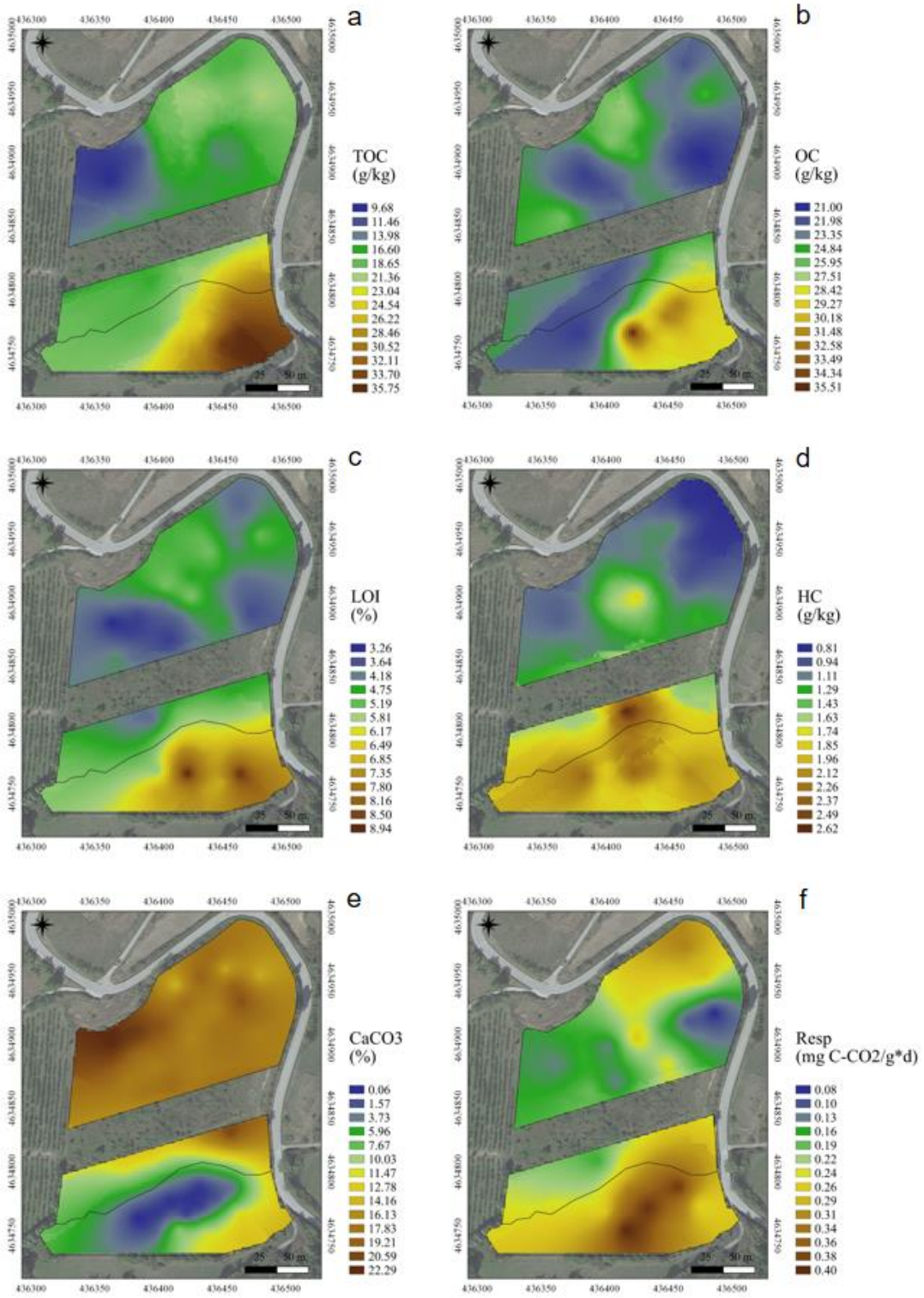


Figure 3.7 Spatial predictive maps of a) total organic carbon ($TOC_{WB}$), b) organic carbon ($OC_{CNS}$), c) loss on ignition (LOI), d) water soluble carbon (HC), e) total calcium carbonate ($CaCO_3$) and f) respiration rate (Resp) content across the study area.

could reflect variations also in soil temperature and moisture which were certainly influential environmental factors controlling soil respiration.

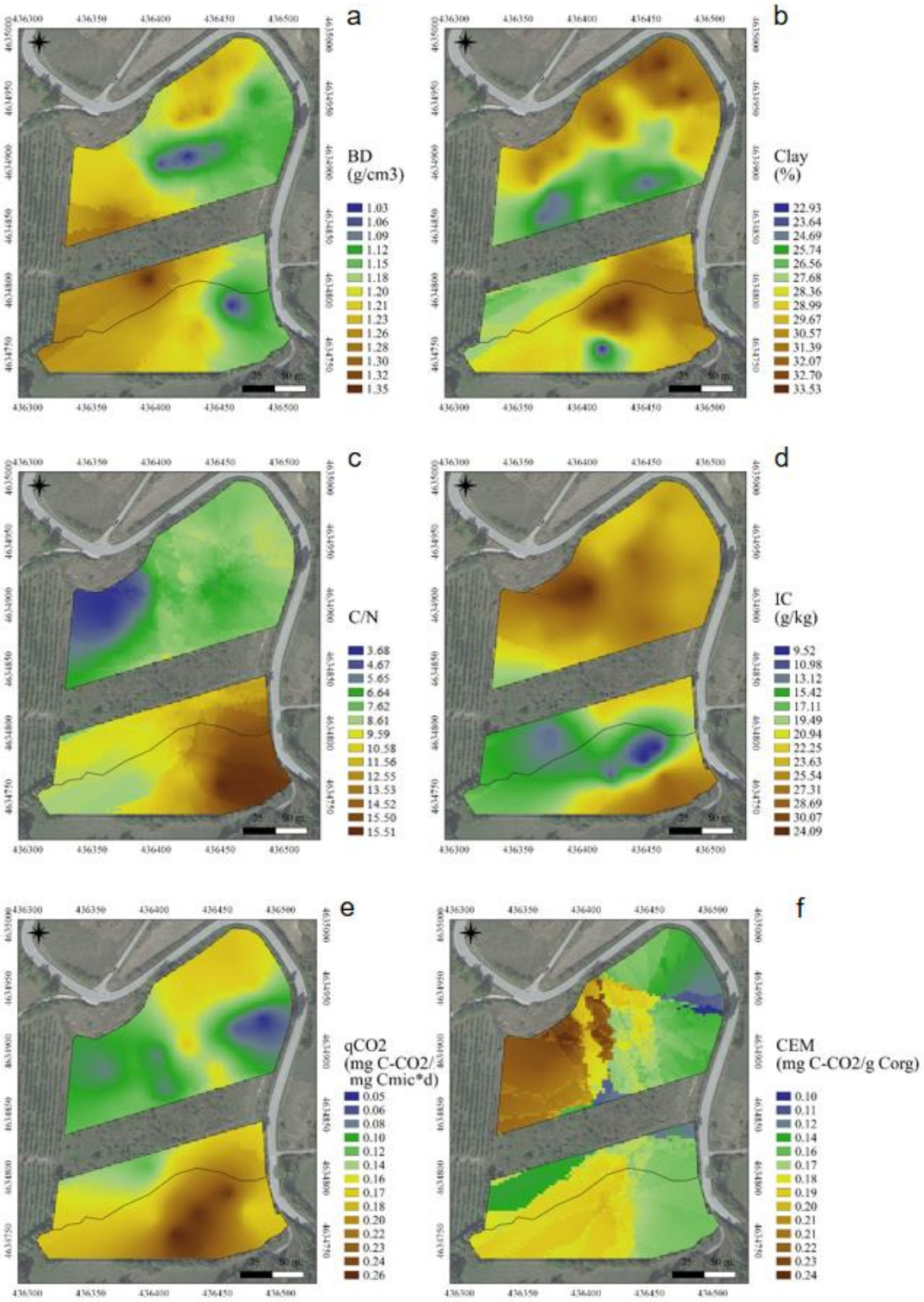


Figure 3.8 Spatial predictive maps of a) bulk density (BD), b) clay content (Clay), c) C/N ratio, d) inorganic carbon ($IC_{CARB}$) content, e) microbial activity ($qCO_2$) and f) coefficient of endogenous mineralization (CEM) across the study area.

Another important consequence of soil parameters distribution was due to correlation between clay content and soil respiration. On the contrary, smaller mineral particles in the soil have the capability of sorb and protect organic matter from degradation, and respiration rate

should have an opposite trend to that of the clay content. However, from what has been found for the studied soils and what Nawar et al. (2016) have said, a small amount of clay was of benefit to microbial soil activity, as they could provide a formation of microaggregate necessary for a stable microbial colonization. In addition particularly evident from the predictive maps of OK, was that the respiration rate of the soil and in general the microbial activity were particularly hindered by the presence of large quantities of calcium carbonate (Figure 3.7e). The studied area has a high variability of total carbonate content, from 0 % in the southern field, to 22 % in the northern one, as resulted also at the time of field sampling, where the soils showed a large difference in effervescence with a field test with HCl. Most likely, those of parent material and pedogenetic processes were the structural factors that had most influenced the spatial variability of carbonates. The high heterogeneity of the parent material that originated the soils of the study area certainly contributed, as evident from the geological map and as already explained in the previous paragraphs. Spatial predictions of inorganic carbon ($IC_{CNS}$) content (Figure 3.8d), obtained by subtracting the organic carbon from the total carbon value (CNS analyzer), were strictly correlated to prediction and spatial trend of total calcium carbonate, demonstrating that calcium carbonate represented most of the inorganic carbon fraction present in the soil.

Lastly, bulk density, as also indicate from semivariogram models (Figure 3.6a and b), showed a short-range autocorrelation, as clay attribute, and they seemed to be strongly correlated each other. In the thematic map (Figure 3.8a), bulk density reached its maximum values (1.2-1.3 g $cm^{-3}$) in correspondence of the areas with the clay maxima (32-33 %), and this was particularly evident in the field to the north, where a lower content of organic matter tended to make the effect of the high clay content prevail. To the southern field, where the accumulation of organic materials was important (35 g $kg^{-1}$ of $TOC_{WB}$), despite the clay reaching quite high contents (33%), the bulk density increase according the typical values of soils with a prevalent sandy fraction (1-1.1 g $cm^{-3}$). This occurrence was totally explained in the presence of the organic material, capable of making the soils less heavy and more aerated, at least in the top soil layer.

Thematic maps in Figure 3.9c and d show, respectively, NDVI and NDWI vegetative indexes for studied fields. Both were computed by using Sentinel-2 remote sensing data in Copernicus Open Access Hub platform, from the tile T33TVG and imagine date corresponding to those of soil sampling, in order to make comparable soil and vegetative sources of data. NDVI describes the vigor level of the crop and was calculated as the ratio between the difference and the sum of the radiations reflected in the near infrared and in the red, i.e. as (NIR-

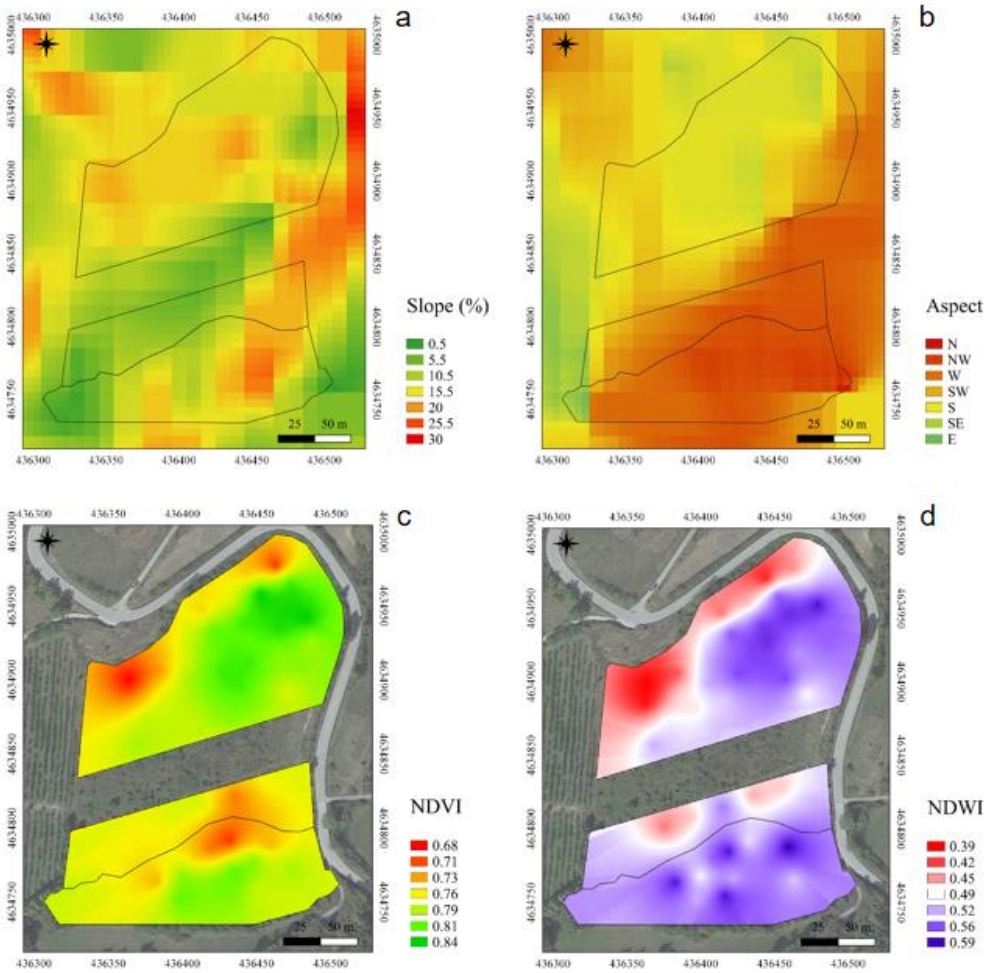


Figure 3.9 a) Slope and b) aspect of the studied fields, derived from DEM with 5 m precision. Spatial predictive maps of c) NDVI and d) NDWI, both vegetative indexes computed with Sentinel-2 remote sensing data.

Red)/(NIR + Red). The interpretation of the absolute value of the NDVI is highly informative, since it allows you to immediately recognize the areas of the field that have development problems. Values of NDVI that ranged from 0.6 to 0.8 (Figure 3.9c) indicated a condition of general high vegetation cover with medium-high vigor for the entire studied area. However, normalizing data in the map through association of the red color with the lowest value and the green one with the highest value, it was possible to highlight the areas of the field most stressed, which coincided with the northernmost part of the study area. The same situation occurred in Figure 3.9d, which shows the NDWI values. NDWI describes the water stress level of the crop and was calculated as the ratio between the difference and the sum of the radiations reflected in the near infrared and in the SWIR, i.e. as (NIR-SWIR)/(NIR + SWIR). The interpretation of the absolute value of the NDWI allowed to immediately recognizing the areas of the field that have water stress problems. Values, ranged from 0.4 to 0.6, indicate a high vegetation cover with no

water stress, for the entire studied area. However, the parts of the field that were most affected by water stress and which have a lower vigor, corresponded to those in which there were the greatest values of calcium carbonates content, demonstrating that the excessive presence of carbonates, typical of the northern part of the study area, negatively affected the vegetation.

## 3.3 Predictions of soil carbon pools using Vis-NIR data

### 3.3.1 Vis-NIR spectra characterization

Soil spectral data were acquired using a spectrophotometer in the UV-Vis-NIR range (200-2500 nm). Prior to spectra acquisition, soil samples were air-dried and finely grounded, to minimum reduce the scattering light effect and homogenize samples. Figure 3.10 shows the raw reflectance spectra of 56 soil samples, for each of which mean spectrum (solid red line) and standard deviation values (shaded regions) were highlighted. In particular, no treated (NT), HCl treated (HClT) and hot water extraction treated (HWT) soil samples showed important differences of the reflectance spectra (Figure 3.10a, b and c, respectively).

Soil spectra showed several large bands of absorbance in the NIR, between 1400 and 2200 nm, in the spectral signatures of all three types of sample treatments. The effects of water absorption near 1400, 1900, and 2200 nm were evident in all raw reflectance and were strengthened in the HWT soil curves (Figure 3.10c). For the raw reflectance, the NT and HWT soil curves had a similar pattern, but with a small difference in the reflectance values, and a moderate shifting of the 1900 nm band up to 1940 nm. (Figure 3.10b). This difference were not linear across the spectrum because the visible part was less affected than the NIR region. When looking at the HWT soil curves, apart from the minor differences over the range of 300–410 nm, the main difference was the shifting of the band in the range 1900-1940 nm. For instance, wavelengths at 1900 and 2200 nm characterized absorptions from a second overtone of carboxyl C=O stretch vibration. However, the 1900-2200 nm was also due the Si-O adsorption vibrations that could overlaps with an absorption at 1940 nm caused by a combination vibration of water in the lattice structure of 2:1 and 2:2 clay mineral (Viscarra Rossel and Hicks 2015). The band at 1400 nm apart of the water was usually also associated with O-H vibration of organic molecules and aliphatic C-H, whilst the absorbance band at 1900 nm was related with amide N-H and O-H. In the band of 2200 nm, there are many organic functional groups such as phenolic O-H, amide N-H, amine H and aliphatic C-H (Fidêncio et al., 2002; Cozzolino and Morón, 2003). Overtones and combinations of fundamental stretching of these bands occurred

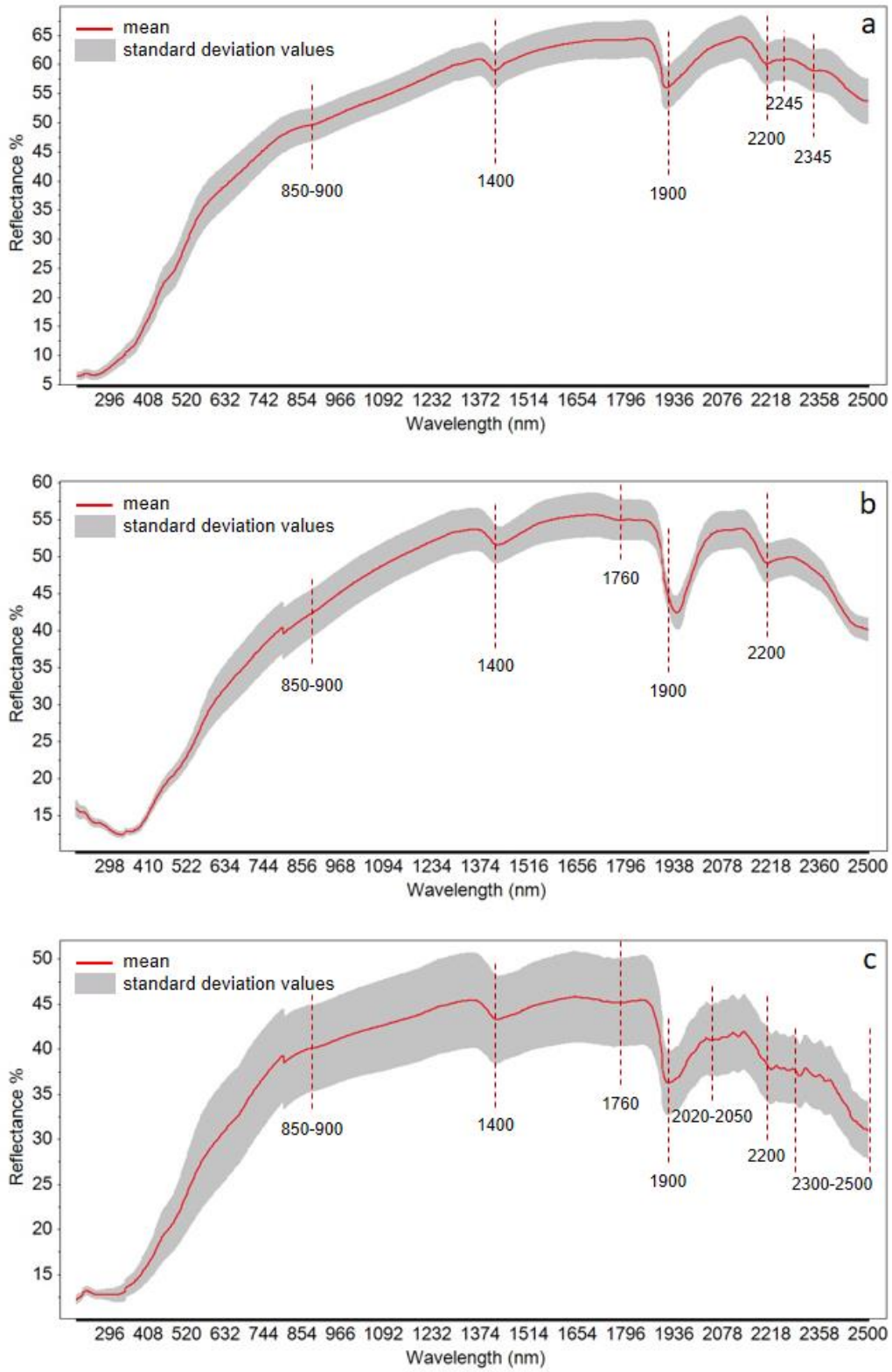


Figure 3.10 The mean soil spectrum and standard deviation values (shaded regions) in the Vis-NIR spectral range of a) no-treated soil samples (NT); b) HCl treated soil samples (HClT) and c) hot water extraction treated (HWT) soil samples.

in the NIR region, and were associated to molecular O-H into the clay minerals and surface water (Baumgardner et al., 1986). As suggested from Asgari et al. (2020), the spectral range

between 350 and 1150 nm (i.e. the visible plus NIR) corresponds to the region with strong SOM influence. Samples containing lower SOM normally showed lower reflectance intensity in the Vis than samples having higher amount because they were darker. In this spectral region (350-1150 nm), absorption features at 850-900 nm could be also related with the presence of organic matter, particularly evident in the spectra curve of NT sample (Figure 3.10a). With reference to the HWT (Figure 3.10c), the higher standard deviation of reflectance values compared to the others (Figure 3.10a and b) was probably due to the high variability of the soluble organic fraction (HWT), compared with the heavy organic fraction ($TOC_{WB}$, $OC_{CNS}$ and LOI). In addition, the high variability of the reflectance values acquired from HWT samples was also attributable to a lower concentration of inorganic carbon and mineral fractions than those contained in soil samples. As shown in Figure 3.10c, spectra of HWT extracts showed several better clear band in the NIR region, consistent with the typical characteristics of the organic functional groups. The band at 1760 nm, which was found by Workman and Weyer (2008) to be assignable to the functional aliphatic groups and aryl C-H functional groups, was particularly evident in Figure 3.10 b and c, and less in the curve in the Figure 3.10 a, where the presence of calcium carbonate in no-treated soils probably overlapped this organic molecules feature. Others important features were found in the range 2020-2050 nm, due to presence of methyl C-H groups (carbohydrate), which distinguished the chemical composition of the non-recalcitrant fractions of organic carbon (Viscarra Rossel and Hicks, 2015). In addition, absorption features between 2300 and 2500 nm (Figure 3.10c) could be attributed to aromatics C-H, phenolics C-OH and aliphatic C-H links (Viscarra Rossel and Hicks, 2015) and not decomposed organic residue, including cellulose, glucan and pectin with different roles in the turnover of SOC.

Spectral curves showed that all these reflections overlapped to signals of non-SOC components in the spectral signatures of NT soil (Figure 3.10a), indicating that the HWT extracts was very useful to evaluate the effect that the labile component of the soil organic matter (HWT) had on the spectral behavior.

From comparison between Figure 3.10a and b, it was possible to evaluate the contribution of inorganic carbon fraction on the spectra features. Firstly, as also suggested by Riefolo et al. (2019), spectrum with the highest total carbonate content showed a greater reflectance than the lower one. Indeed, spectra of analyzed soils that did not contain calcium carbonate (previously HCl treated and oven-dried) showed an average lower reflectance of about 8.6% compared to NT curves soils (Figure 3.10a). Furthermore, as observed by several authors (Ben-Dor and Banin, 1990; Riefolo et al., 2019; Asgari et al., 2020), in the NIR spectral range the strong bands that appear at around 2245 and 2345 nm, were due to C-O stretching mode in $CaCO_3$ molecules.

These findings indicated the presence of carbonates and identified the spectral range of 2000-2500 nm as highly informative for their quantification.

### 3.3.2 Soil carbon pools affecting Vis-NIR spectra

Pearson's correlation coefficient (r) at 0.05 *p*-value was used to investigate the correlations of the soil organic and inorganic carbon fractions with the spectral raw reflectance values across the measured Vis-NIR spectral range (Figure 3.11). Generally, all soil properties significantly (*p*-value <0.05) affected the spectral reflectance response of the soil samples. Total carbon ($TC_{CNS}$) obtained by using CNS analyzer, is the carbon fraction of the soil that was least correlated to the spectral information acquired, particularly in the Vis range, where correlations were not significant (r = 0.05-0.2). In the NIR range, $TC_{CNS}$ exhibited a poorly positive correlation across the spectrum.

As a first indicative result, it was possible to note how the fractions of inorganic carbon, expressed by calcium carbonate content ($CaCO_3$) and inorganic C ($IC_{CNS}$) obtained by $TC_{CNS} - OC_{CNS}$ (CNS analyzer), were positively correlated to soil spectra. Whereas, organic fractions, expressed as total organic carbon ($TOC_{WB}$) and hot water carbon (HC) by wet oxidation, and $OC_{CNS}$ and LOI by dry combustion methods, negatively correlated soil spectra. This finding corroborated the previous results obtained from analysis of soil spectra characterization, according to which the greater content of carbonate increased the reflectance of soils, while the greater content of organic substance decreased soils reflectance, in particular for Vis range.

Among positive correlations, the highest ones were found for $CaCO_3$, followed by $IC_{CNS}$. The effect of $CaCO_3$ content on spectral reflectance was more intensive in the NIR region with reaching its higher values in the range 0.80- 0.83 between 1400 and 2500 nm wavelength. These results are consistent with those of Mammadov et al. (2020). The patterns observed for $CaCO_3$ were similar for $IC_{CNS}$ content, although this showed a lower correlation with spectral information than the carbonate content. This was probably due to the additional presence of inert carbon in the $IC_{CNS}$ fraction compared to that of $CaCO_3$, exclusive to calcium carbonate content. Then $IC_{CNS}$ generated a less precise response on the spectral signatures of soils due to spectral overlapping of inorganic carbon components that differ in their chemical composition.

Regarding the carbon fractions negatively correlated to spectral reflectance, the highest and lowest correlations were typical to LOI and $TOC_{WB}$ fractions, respectively. In general, all the measured organic fractions ($TOC_{WB}$, $OC_{CNS}$, LOI and HC) showed a greater negative correlation in the Vis range rather than in the NIR, demonstrating that the Vis was the spectral

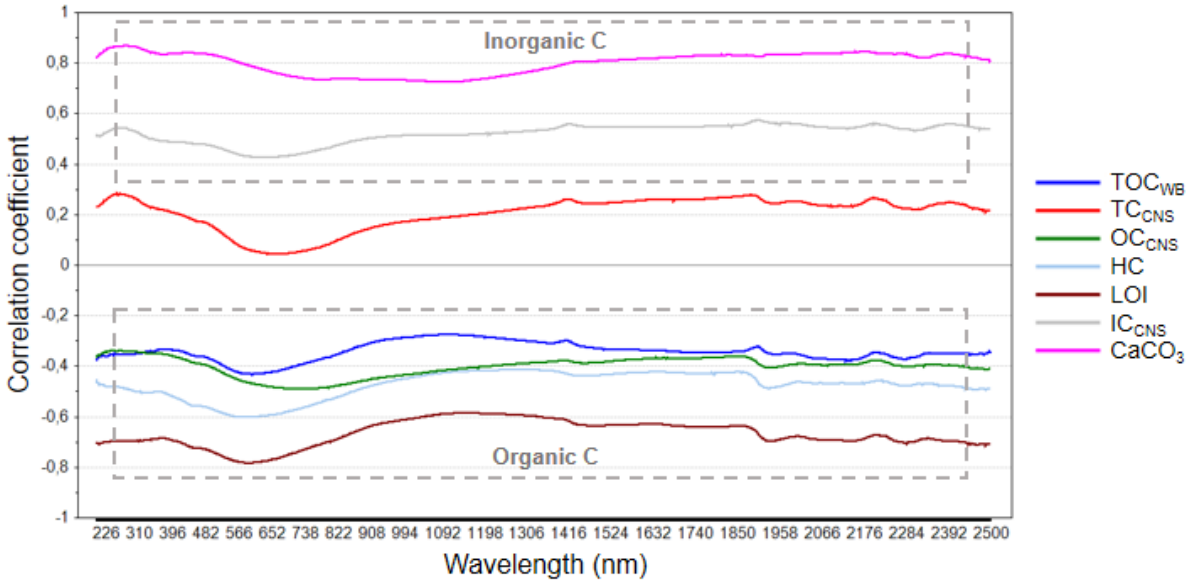


Figure 3.11 Pearson's correlation coefficients between spectral reflectance at each wavelength (200-2500 nm) and soil carbon pools.

range for which the SOM values had a lower reflectance. The correlations were especially higher (-0.78 for LOI; -0.6 for HC; -0.5 for $OC_{CNS}$ and -0.45 for $TOC_{WB}$) at wavelengths between 540 nm and 730 nm as observed by previous studies (Henderson et al., 1992; Bartholomeus et al., 2008; Nocita et al., 2013), who found highest correlation at 600 nm wavelength. Summarily, this research showed variations and typical patterns in reflectance spectra corresponding to the variations in soil carbon pools.

In addition, to deepen the knowledge of the link between spectra and soil carbon fractions as described in the previous paragraph (descriptive statistics of soil properties), principal component analysis (PCA) was also performed using the soil spectra wavelength as continuous variables for construction of the new system of components. The first interesting implication resulting from this further processing was the better percentage of total variance of dataset explained by the first two components. As shown in Figure 3.12, PC1 accounted for 94% and PC2 for 4%, for 98% of the entire variance of the dataset. The great difference in the percentages of variance explained between soil property variables (referred to Figure 3.4) and spectral variables surely lied in the first place from the difference between the numbers of variables used in the two PCAs (35 soil properties compared to 1150 reflectance points).

However, as shown in Figure 3.13, and differently from what happened in the previous PCA (referred to Figure 3.4), all the reflectance values for the spectra of the soils were considered as important for the construction of the new system of components capable of describing the dataset. Indeed, PC1, represented by blue line in the figure, was entirely comprised between the

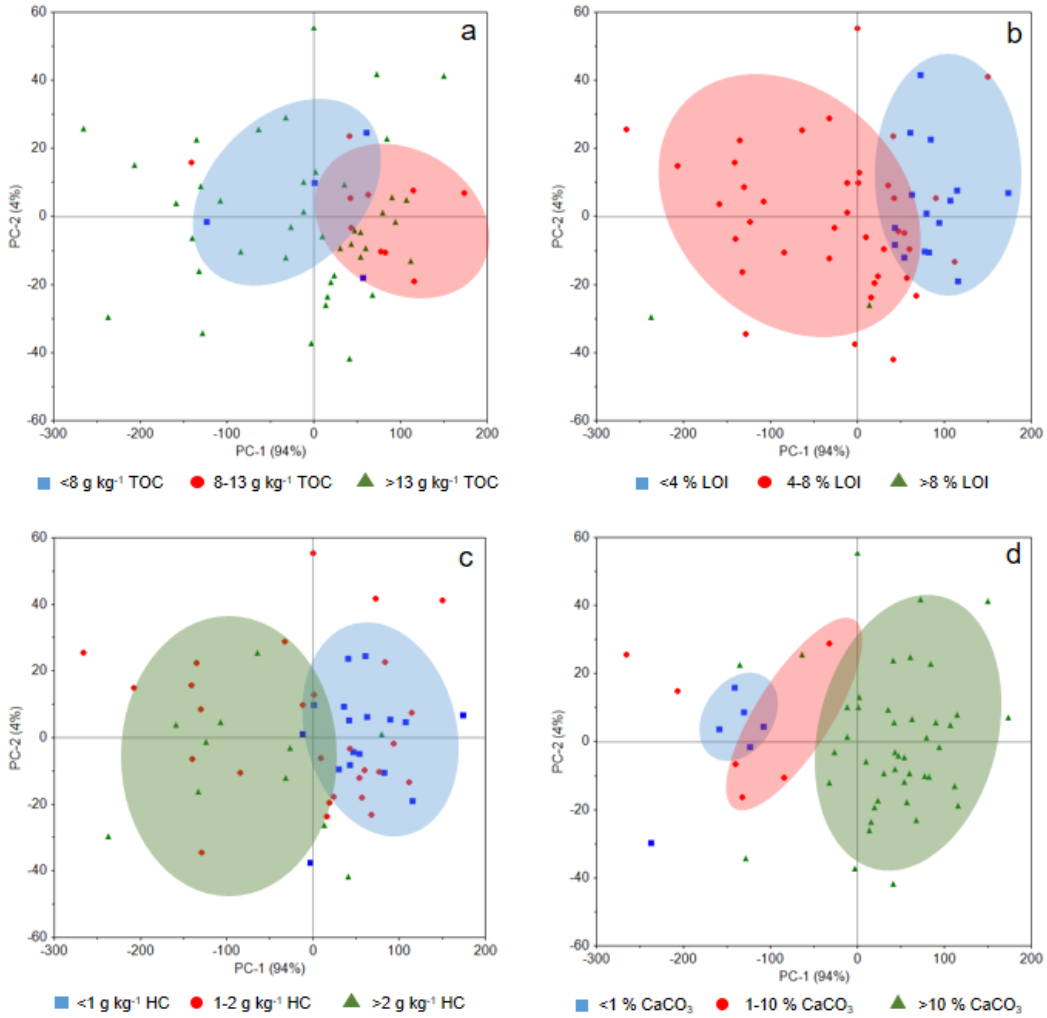


Figure 3.12 Score plot resulting from PCA performed using soil spectra wavelength as X-loadings. Groupings refer to clusters of samples according to scaling soil parameters: a) organic carbon content, expressed as $TOC_{WB}$; b) LOI parameter; c) extractable hot water organic carbon (HC) and d) calcium carbonate contents ($CaCO_3$).

two dashes red lines, which represented respectively the 50 and 100% of contribution in the model. This result highlighted the absolute but not exclusive importance of using row Vis-NIR reflectance spectra for a first evaluation of a set of soil samples.

In addition, the score plot resulting from this second application of PCA was used for demonstrating the presence of any clusters of samples, according to soil properties and using only soil spectra as variables. Properties about carbon pool contents in the soil were chosen: inorganic ($CaCO_3$), expressed as percentage of total calcium carbonate, and organic, expressed by both total organic carbon ($TOC_{WB}$ from Walkley-Black method) and LOI. Classification of values of chosen properties was made according to VV.AA. (2006). From Figure 3.12d, Vis-NIR spectra were able to highlight very well the different calcium carbonate contents in the studied soils, from weakly calcareous (blue squares) to medium (red dots) and strongly

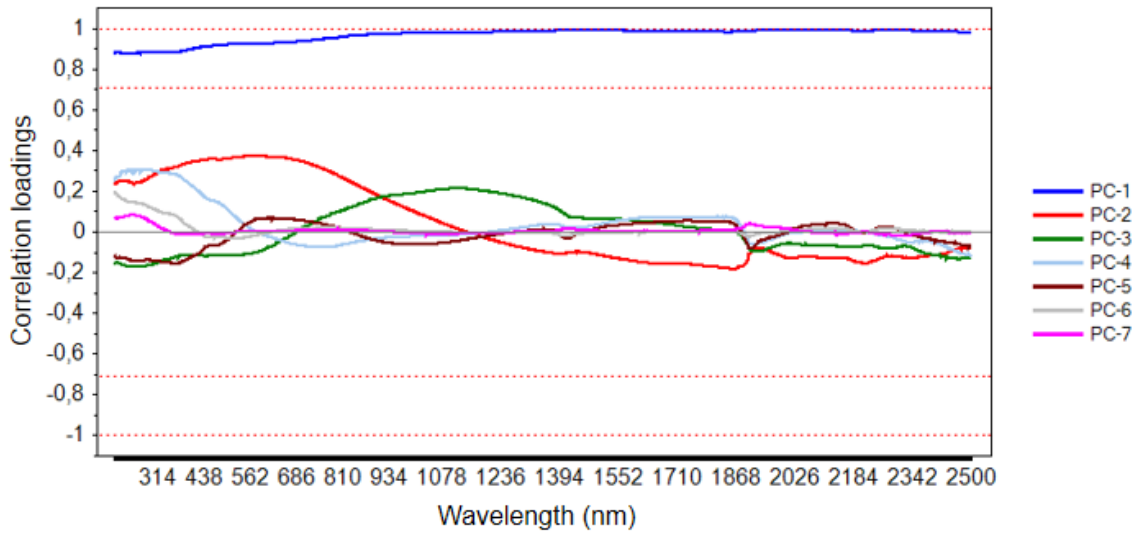


Figure 3.13 Correlation loadings line plot resulting from PCA performed using soil spectra wavelength as X-loadings.

calcareous (green triangles). The same result was found for the LOI parameter (Figure 3.12b): excellent discrimination between samples with good LOI content (<4%, blue squares) and rich content (4-8%, red dots) was shown; the highest values of LOI (>8%, green triangles) were not well represented, as they were present only in the number of two samples. Instead, for $TOC_{WB}$ parameter, Vis-NIR spectra were able to separate only soil samples with the lowest organic carbon contents (Figure 3.12a); blue squares highlight values lower than 8 g $kg^{-1}$ of $TOC_{WB}$, red dots values from 8 to 13 g $kg^{-1}$ of $TOC_{WB}$. For HC parameter, which represents the soil carbon pool more labile and active and more easily degradable by microorganisms, soil spectra were able to separate only the lower HC values (blue squares; HC <1 g $kg^{-1}$) from the higher ones (green triangles; HC >2 g $kg^{-1}$). This suggested that the used scaling parameter did not allow the intermediate values of HC to be well identified, which presented a rather generic and confusing placement in the score plot (Figure 3.12c). Further, both $TOC_{WB}$ and HC organic fractions that were obtained by wet combustion procedures and then subjected to incomplete oxidation were less clearly highlighted by the spectral data than the LOI parameter. Since the LOI was measured by a dry combustion technique that generated an overestimation of the organic substance mainly due to the presence of carbonates and clay minerals (Ben-Dor and Banin, 1989), it was less affected by the interference and overlapping of Vis-NIR signal due to the non-SOC components (Chatterjee et al., 2009). This leaded to a better subdivision of the samples with different LOI content as a function of the spectral data. However, it was possible to state that all the values taken into consideration in the PCA were distinct according to PC1 (scaling according x-axis), as even suggested by the correlation loadings plot in Figure 3.13.

Anyway, it was possible to state that Vis-NIR reflectance spectra were suitable for both a first check of soil samples, and also to discern for their soil carbon pools, either inorganic or organic. The way in which the spectra are able to evaluate each carbon pool in the soil will be discussed later.

### 3.3.3 Performances of carbon pools predictions

The predictions of soil carbon pools in the spectral Vis-NIR range were performed randomly splitting the studied dataset in calibration (70%) and validation (30%) subsets. To reduce the spectral noise and optimize the spectral data, different kinds of preprocessing transformations were applied (see the specific preprocessing transformation techniques and related abbreviations in Table 3.4). Three different regression models were chosen, partial least square (PLS), support vector machine (SVM) and boosted tree (BT). Prior to regression models application, raw reflectance spectra were treated with 5 different preprocessing, in order to enhance the spectral features and fit the best relationship with the different organic carbon pools of the soil. The adopted preprocessing were: Savitzky-Golay filter and absorbance transformation (SG+Abs), first and second derivative transformations (1st Derivative and 2nd Derivative), standard normal variate (SNV) and multiplicative scatter correction (MSC) transformations. Preprocessing were also compared in all regression methods with performances of predictions obtained from raw reflectance spectral data (R). Tables 3.4 and 3.5 summarizes the results of predictions (calibration and validation respectively), whose performances were evaluated by means of coefficient of determination ($R^2$), root mean squared error (RMSE) and residual prediction deviation (RPD).

In general, the application of regression models and preprocessing has obtained good results in calibration (Table 3.4), although there were differences of performances among models, preprocessing and even when considering soil organic and inorganic carbon pools. When considering the model quality in validation (Table 3.5), predictions provided the best results, with coefficient of determination ($R^2$) varying from 0.90 for the PLS model to 0.93 for the $CaCO_3$ content. Usually, preprocessing transformations of spectral data improve the accuracy of regression models. Some studies have reported improvements of the regression models by using first and second derivatives (Dunn et al., 2002), normalization of the data, and scatter corrections, while others found better results with untransformed reflectance data (McCarty et al., 2002). Figure 3.14 and 3.15 display values of coefficient of determination $R^2$ resulting from validation of the predictions about organic and inorganic carbon fractions, respectively. Firstly, it was possible to note that for the majority of cases, the application of preprocessing

transformations has led to improving of predictions, compared to results of regressions obtained using directly raw reflectance spectra, without any pre-treatment. This finding was in line with different authors (Igne et al., 2010; Buddenbaum and Steffens, 2012; Zhang et al., 2020; Davari et al., 2021), who found improving of the prediction accuracy with pre-treated soil spectra.

Among the four organic carbon fractions ($TOC_{WB}$, $OC_{CNS}$, LOI and HC; Figure 3.14) and the two inorganic carbon ($IC_{CNS}$ and $CaCO_3$; Figure 3.15), the best performances of predictions in validation were obtained from LOI and $CaCO_3$ content parameters (Table 3.5). The prediction of LOI had achieved excellent values of coefficient of determination, corresponding to $R^2$ = 0.84 for from the PLS regression model with 1st derivative transformation (RMSE = 0.67, RPD = 2.36) and SVM regression model with 2nd derivative transformation (RMSE = 0.69, RPD = 1.79), and $R^2$ = 0.85 for SVM model with 1st derivative transformation (RMSE = 0.66, RPD = 2.02). Values of RPD >2.0 were considered accurate predictive models; RPD values between 1.4 and 2.0 indicated fair models that could be improved by more accurate predictive techniques; RPD values <1.4 indicated poor predictive capacity (Chang et al., 2001; Chang and Laird, 2002). Indeed, as shown in Figure 3.14c, PLS and SVM turned out to be the regression models that best fit the values of LOI, compared to BT regression, for which derivatives, SNV and MSC preprocessing transformations led to degrading the prediction performance. Moreover, the combination between preprocessing method and regression model that generated the best accuracy of prediction was 2nd derivative/PLS.

Considering predictions from no-pretreated spectra, the prediction accuracy improved in terms of $R^2$ from 0.22 to 0.79 for $TOC_{WB}$, from 0.18 to 0.80 for $OC_{CNS}$ and from 0.45 to 0.90 for HC fraction, with correspondig RMSE and RPD values of 0.43 and 1.93, 2.87 and 2.83, and 0.02 and 3.05 for $TOC_{WB}$, $OC_{CNS}$ and HC, respectively. However, as possible to see from Figure 3.14, 2nd derivative transformation was also able to improve the estimation accuracy of the SVM regression, consistently of what found by Gao et al. (2014), for which the good performance achieved by the SVM model can be clarified by its capability to model the nonlinearity among the reflectance data and soil organic carbon content. These results match those from Dotto et al. (2018), who found the spectral derivative preprocessing with the SVM method the best combination for the prediction of SOC content. Performances of prediction of HC fraction (Figure 3.14d) were also good, suggesting that the spectral analysis was able to successfully predict even the labile fraction of soil organic carbon, which represents the organic part most exposed to degradative and erosive phenomena. Lastly, BT regression on organic carbon pools was the model that achieved the lowest results in terms of accuracy, excluding

Table 3.4
Performances of calibration of soil carbon pools ($TC_{CNS}$, $TOC_{WB}$, $OC_{CNS}$, HC, LOI, $IC_{CNS}$ and $CaCO_3$) predictive models in Vis-NIR spectra with the corresponding preprocessing transformation.

| Pre-proc | Model | $TC_{CNS}$ | | | $TOC_{WB}$ | | | $OC_{CNS}$ | | | HC | | | LOI | | | $IC_{CNS}$ | | | $CaCO_3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD |
| Raw spectra | PLS | 0.65 | 0.44 | 1.36 | 0.44 | 0.69 | 0.90 | 0.18 | 5.84 | 0.47 | 0.43 | 0.06 | 0.87 | 0.83 | 0.70 | 2.21 | 0.71 | 4.09 | 1.58 | 0.94 | 1.65 | 3.85 |
| | SVM | 0.34 | 0.65 | 0.48 | 0.39 | 0.58 | 0.55 | 0.02 | 4.36 | 0.23 | 0.29 | 0.07 | 0.48 | 0.59 | 0.96 | 1.08 | 0.42 | 5.35 | 0.91 | 0.89 | 2.08 | 2.75 |
| | BT | 0.06 | 0.69 | 0.57 | 0.90 | 0.29 | 2.65 | 0.07 | 7.02 | 0.27 | 0.86 | 0.03 | 2.31 | 0.89 | 0.62 | 2.64 | 0.51 | 5.18 | 0.88 | 0.93 | 1.61 | 3.72 |
| SG+Abs | PLS | 0.66 | 0.43 | 1.42 | 0.45 | 0.69 | 0.91 | 0.36 | 5.17 | 0.75 | 0.43 | 0.06 | 0.87 | 0.81 | 0.75 | 2.06 | 0.63 | 4.60 | 1.33 | 0.96 | 1.26 | 5.07 |
| | SVM | 0.35 | 0.65 | 0.50 | 0.39 | 0.58 | 0.55 | 0.02 | 4.38 | 0.24 | 0.29 | 0.07 | 0.48 | 0.58 | 0.97 | 1.04 | 0.43 | 5.31 | 0.95 | 0.88 | 2.11 | 2.71 |
| | BT | 0.02 | 0.72 | 0.55 | 0.81 | 0.40 | 1.76 | 0.15 | 6.74 | 0.30 | 0.83 | 0.03 | 2.09 | 0.89 | 0.63 | 2.64 | 0.53 | 5.05 | 0.92 | 0.92 | 1.63 | 3.68 |
| 1st Derivative | PLS | 0.83 | 0.30 | 2.25 | 0.51 | 0.65 | 1.02 | 0.24 | 5.61 | 0.57 | 0.42 | 0.06 | 0.86 | 0.85 | 0.65 | 2.45 | 0.80 | 3.36 | 2.05 | 0.95 | 1.47 | 4.34 |
| | SVM | 0.86 | 0.31 | 2.11 | 0.82 | 0.31 | 1.99 | 0.59 | 2.78 | 0.86 | 0.69 | 0.04 | 1.08 | 0.94 | 0.37 | 3.62 | 0.86 | 2.66 | 2.21 | 0.98 | 0.91 | 6.71 |
| | BT | 0.09 | 0.75 | 0.79 | 0.84 | 0.37 | 1.95 | 0.21 | 6.47 | 0.52 | 0.48 | 0.05 | 1.16 | 0.51 | 1.33 | 1.39 | 0.94 | 1.76 | 3.66 | 0.46 | 4.36 | 1.34 |
| 2nd Derivative | PLS | 0.96 | 0.15 | 4.87 | 0.52 | 0.64 | 1.04 | 0.27 | 5.50 | 0.62 | 0.61 | 0.05 | 1.26 | 0.92 | 0.49 | 3.33 | 0.28 | 6.44 | 0.63 | 0.97 | 1.20 | 5.30 |
| | SVM | 0.95 | 0.18 | 3.72 | 0.95 | 0.17 | 3.93 | 0.87 | 1.57 | 2.07 | 0.94 | 0.02 | 3.61 | 0.94 | 0.35 | 3.67 | 0.95 | 1.64 | 3.73 | 0.98 | 0.93 | 6.10 |
| | BT | 0.06 | 0.69 | 0.75 | 0.87 | 0.34 | 2.43 | 0.13 | 7.75 | 0.36 | 0.91 | 0.02 | 3.03 | 0.19 | 1.71 | 0.87 | 0.87 | 2.71 | 2.10 | 0.62 | 3.65 | 1.61 |
| SNV | PLS | 0.73 | 0.39 | 1.65 | 0.43 | 0.70 | 0.88 | 0.24 | 5.63 | 0.56 | 0.42 | 0.06 | 0.85 | 0.79 | 0.77 | 1.97 | 0.79 | 3.50 | 1.94 | 0.96 | 1.25 | 5.08 |
| | SVM | 0.55 | 0.54 | 0.88 | 0.55 | 0.50 | 0.83 | 0.17 | 3.95 | 0.44 | 0.39 | 0.06 | 0.66 | 0.83 | 0.61 | 1.98 | 0.61 | 4.40 | 0.92 | 0.90 | 1.93 | 3.19 |
| | BT | 0.76 | 0.35 | 1.65 | 0.83 | 0.38 | 2.05 | 0.54 | 4.93 | 0.80 | 0.94 | 0.02 | 3.53 | 0.56 | 1.26 | 1.39 | 0.92 | 2.08 | 3.09 | 0.76 | 2.92 | 1.89 |
| MSC | PLS | 0.73 | 0.38 | 1.68 | 0.43 | 0.70 | 0.88 | 0.24 | 5.63 | 0.56 | 0.42 | 0.06 | 0.85 | 0.79 | 0.77 | 1.97 | 0.73 | 3.94 | 1.66 | 0.95 | 1.43 | 4.42 |
| | SVM | 0.56 | 0.54 | 0.88 | 0.55 | 0.50 | 0.84 | 0.17 | 3.94 | 0.45 | 0.39 | 0.06 | 0.66 | 0.83 | 0.61 | 1.99 | 0.61 | 4.40 | 0.92 | 0.90 | 1.93 | 3.19 |
| | BT | 0.77 | 0.34 | 1.61 | 0.87 | 0.33 | 2.38 | 0.33 | 5.96 | 0.57 | 0.49 | 0.05 | 1.06 | 0.55 | 1.27 | 1.38 | 0.72 | 3.90 | 1.18 | 0.99 | 0.32 | 18.68 |

Note Preprocessing: raw reflectance spectra (Raw spectra); Savitzky-Golay filter and absorbance (SG+Abs); first derivative (1st Derivative); second derivative (2nd Derivative); standard normal variate (SNV) and multiplicative scatter correction (MSC) transformations.
Models: partial least square (PLS); support vector machine (SVM) and boosted tree (BT) regressions.

Table 3.5
Performances of validation of soil carbon pools ($TC_{CNS}$, $TOC_{WB}$, $OC_{CNS}$, HC, LOI, $IC_{CNS}$ and $CaCO_3$) predictive models in Vis-NIR spectra with the corresponding preprocessing transformation.

| Pre-proc | Model | $TC_{CNS}$ | | | $TOC_{WB}$ | | | $OC_{CNS}$ | | | HC | | | LOI | | | $IC_{CNS}$ | | | $CaCO_3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD | $R^2$ | RMSE | RPD |
| Raw spectra | PLS | 0.23 | 0.65 | 0.56 | 0.22 | 0.82 | 0.54 | 0.18 | 5.85 | 0.47 | 0.45 | 0.06 | 0.90 | 0.63 | 1.03 | 1.32 | 0.34 | 6.16 | 0.73 | 0.89 | 2.13 | 2.91 |
| | SVM | 0.29 | 0.63 | 0.48 | 0.26 | 0.80 | 0.42 | 0.14 | 6.88 | 0.15 | 0.29 | 0.07 | 0.52 | 0.63 | 1.04 | 1.05 | 0.48 | 5.50 | 0.85 | 0.89 | 2.18 | 2.65 |
| | BT | 0.21 | 0.66 | 0.76 | 0.62 | 0.57 | 1.24 | 0.00 | 6.44 | 0.37 | 0.48 | 0.06 | 1.03 | 0.68 | 0.96 | 1.56 | 0.45 | 5.66 | 0.86 | 0.86 | 2.47 | 2.52 |
| SG+Abs | PLS | 0.25 | 0.65 | 0.58 | 0.22 | 0.82 | 0.54 | 0.27 | 5.51 | 0.61 | 0.44 | 0.06 | 0.90 | 0.64 | 1.02 | 1.33 | 0.36 | 6.09 | 0.75 | 0.90 | 2.01 | 3.11 |
| | SVM | 0.30 | 0.62 | 0.49 | 0.25 | 0.80 | 0.41 | 0.16 | 6.95 | 0.15 | 0.30 | 0.07 | 0.53 | 0.62 | 1.05 | 1.01 | 0.49 | 5.44 | 0.88 | 0.89 | 2.13 | 2.73 |
| | BT | 0.15 | 0.68 | 0.64 | 0.49 | 0.66 | 1.04 | 0.11 | 6.10 | 0.39 | 0.47 | 0.06 | 1.00 | 0.65 | 1.00 | 1.65 | 0.46 | 5.58 | 0.90 | 0.80 | 2.92 | 2.33 |
| 1st Derivative | PLS | 0.44 | 0.55 | 0.90 | 0.48 | 0.67 | 0.96 | 0.42 | 4.89 | 0.87 | 0.55 | 0.05 | 1.12 | 0.84 | 0.67 | 2.36 | 0.50 | 5.40 | 1.00 | 0.93 | 1.67 | 3.79 |
| | SVM | 0.76 | 0.37 | 1.56 | 0.52 | 0.64 | 0.89 | 0.19 | 5.82 | 0.37 | 0.62 | 0.05 | 1.01 | 0.85 | 0.66 | 2.02 | 0.57 | 5.00 | 1.08 | 0.94 | 1.55 | 4.05 |
| | BT | 0.56 | 0.82 | 0.80 | 0.69 | 0.51 | 1.26 | 0.12 | 6.03 | 0.55 | 0.26 | 0.07 | 0.85 | 0.19 | 1.53 | 1.24 | 0.68 | 4.30 | 1.64 | 0.15 | 5.98 | 0.94 |
| 2nd Derivative | PLS | 0.76 | 0.36 | 1.80 | 0.79 | 0.43 | 1.93 | 0.80 | 2.87 | 2.03 | 0.90 | 0.02 | 3.05 | 0.68 | 0.96 | 1.51 | 0.12 | 7.11 | 0.55 | 0.77 | 3.09 | 1.84 |
| | SVM | 0.83 | 0.31 | 1.90 | 0.52 | 0.64 | 0.93 | 0.40 | 4.99 | 0.57 | 0.84 | 0.03 | 2.02 | 0.84 | 0.69 | 1.79 | 0.58 | 4.90 | 1.11 | 0.89 | 2.17 | 2.52 |
| | BT | 0.13 | 0.69 | 0.60 | 0.65 | 0.55 | 1.30 | 0.25 | 7.21 | 0.41 | 0.41 | 0.06 | 1.16 | 0.01 | 1.69 | 0.99 | 0.54 | 5.18 | 0.95 | 0.28 | 5.52 | 1.38 |
| SNV | PLS | 0.16 | 0.68 | 0.45 | 0.39 | 0.73 | 0.80 | 0.31 | 5.35 | 0.68 | 0.41 | 0.06 | 0.84 | 0.82 | 0.72 | 2.14 | 0.29 | 6.42 | 0.64 | 0.80 | 2.92 | 2.01 |
| | SVM | 0.42 | 0.57 | 0.79 | 0.34 | 0.75 | 0.52 | 0.02 | 6.51 | 0.26 | 0.42 | 0.06 | 0.76 | 0.72 | 0.89 | 1.39 | 0.33 | 6.20 | 0.63 | 0.87 | 2.31 | 2.73 |
| | BT | 0.30 | 0.62 | 0.95 | 0.41 | 0.71 | 1.15 | 0.38 | 5.07 | 0.73 | 0.54 | 0.05 | 1.26 | 0.19 | 1.53 | 1.19 | 0.53 | 5.19 | 1.35 | 0.49 | 4.64 | 1.30 |
| MSC | PLS | 0.23 | 0.65 | 0.56 | 0.39 | 0.73 | 0.80 | 0.27 | 5.51 | 0.61 | 0.41 | 0.06 | 0.84 | 0.81 | 0.73 | 2.11 | 0.29 | 6.39 | 0.65 | 0.80 | 2.90 | 2.02 |
| | SVM | 0.42 | 0.57 | 0.79 | 0.34 | 0.75 | 0.52 | 0.02 | 6.51 | 0.26 | 0.42 | 0.06 | 0.76 | 0.72 | 0.90 | 1.38 | 0.33 | 6.20 | 0.63 | 0.87 | 2.31 | 2.74 |
| | BT | 0.33 | 0.61 | 1.03 | 0.42 | 0.71 | 1.17 | 0.19 | 5.79 | 0.56 | 0.31 | 0.06 | 0.86 | 0.18 | 1.53 | 1.19 | 0.42 | 5.81 | 0.91 | 0.65 | 3.82 | 1.66 |

Note Preprocessing: raw reflectance spectra (Raw spectra); Savitzky-Golay filter and absorbance (SG+Abs); first derivative (1st Derivative); second derivative (2nd Derivative); standard normal variate (SNV) and multiplicative scatter correction (MSC) transformations.
Models: partial least square (PLS); support vector machine (SVM) and boosted tree (BT) regressions.

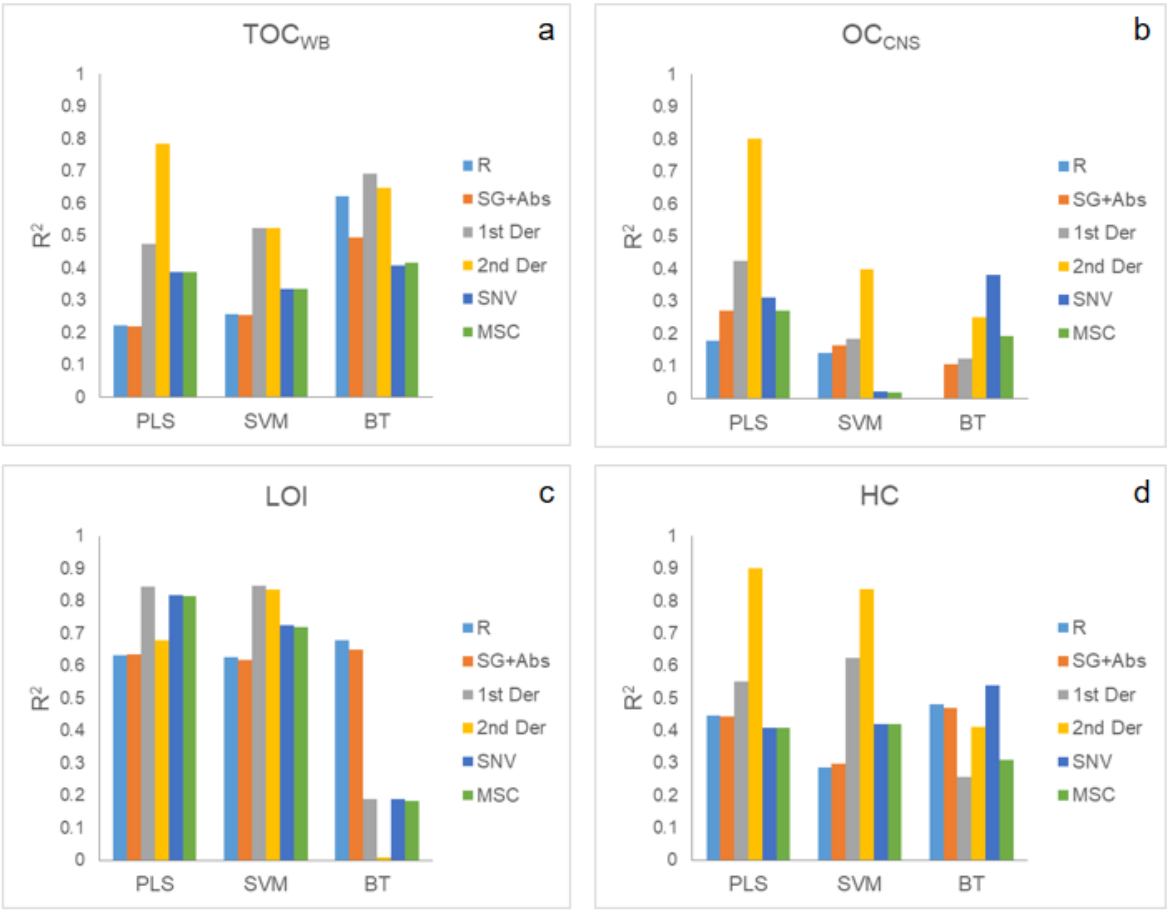


Figure 3.14 Coefficient of determination ($R^2$) in validation of PLS, SVM and BT regressions for soil organic carbon fractions a) $TOC_{WB}$, b) $OC_{CNS}$, c) LOI and d) HC) under different preprocessing.

$TOC_{WB}$ parameter, for which application of this method achieved very good performances starting from derivative spectra ($R^2$ = 0.69, RMSE = 0.51 and RPD = 1.26, in 1st derivative; and $R^2$ = 0.65, RMSE = 0.55 and RPD = 1.30, in 2nd derivative).

Figure 3.15 shows the performance of predictions in terms of coefficient of determination $R^2$ about total and inorganic soil carbon pools. In particular, for total carbon fraction ($TC_{CNS}$; Figure 3.15c) accuracy of predictions were lower than the others analyzed parameters. This finding was expected, given the low correlation between the measured total carbon values and the spectral reflectance across the entire Vis-NIR range, as demonstrated in Figure 3.11 previously discussed. In general, fitting for $TC_{CNS}$ was higher in SVM regression than PLS and BT, with Savitzky-Golay filter and absorbance transformation preprocessing that was performing as well as raw reflectance spectra. However, in line with the results obtained for organic carbon fractions ($TOC_{WB}$, $OC_{CNS}$, LOI and HC), the best preprocessing method that considerably increased the estimation accuracy was the second derivative transformation, so the values of $R^2$, RMSE and RPD were respectively 0.76, 0.36 and 1.80 when combined with

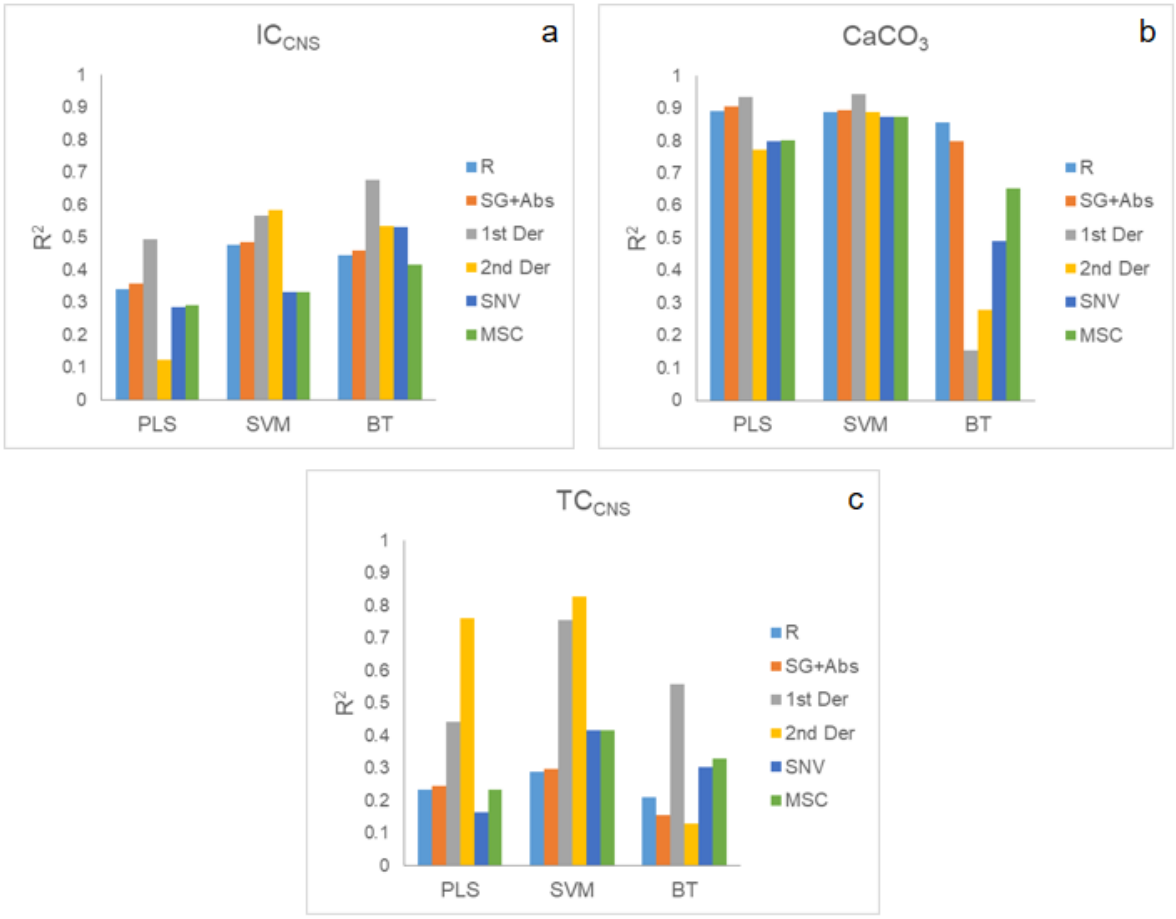


Figure 3.15 Coefficient of determination ($R^2$) in validation of PLS, SVM and BT regressions for soil inorganic carbon fractions a) $IC_{CNS}$ and b) $CaCO_3$ content, and c) total carbon content ($TC_{CNS}$) under different preprocessing.

PLS, and 0.83, 0.31 and 1.90 in combination with SVM regression (Table 3.5).

Prediction performances of soil inorganic carbon fractions, expressed as $IC_{CNS}$ ($TC_{CNS}$ - $OC_{CNS}$) by CNS analyzer and calcium carbonate content ($CaCO_3$) by gas-volumetric method, are shown in Figure 3.15a and b, respectively. As shown, they had opposing capacities to be predicted in regression models by means of Vis-NIR spectra. PLS and SVM regression models were more suitable for $CaCO_3$ parameter than $IC_{CNS}$, whose predictions showed more performing results when applied the BT regression method. as shown in the Figure 3.15a and the Table 3.5, $IC_{CNS}$ gained the best performance of prediction by combination of 1st derivative preprocessing transformation and BT regression model, with $R^2 = 0.68$, RMSE = 4.30 and RPD = 1.64.

On the contrary, the $CaCO_3$ parameter, which achieved the best prediction results among all the carbon pools analyzed, showed the finest results ($R^2 = 0.93$ and 0.94) with the combinations of 1st derivative/PLS and 1st derivative SVM, respectively, but excellent results were also achieved applying all the used preprocessing transformations in PLS and SVM, comprises raw

reflectance spectra. No treated spectra for $CaCO_3$ have obtained results of $R^2$ = 0.89, RMSE = 2.13 and RPD = 2.91 in PLS, $R^2$ = 0.89, RMSE = 2.18 and RPD = 2.65 in SVM, and $R^2$ = 0.86, RMSE = 2.47 and RPD = 2.52 in BT regression models, as displayed in Table 3.5. These findings were in agreement with Miloš and Bensa (2018), who found slightly better results with SVM than PLS regression ($R^2$ = 0.88 compared to $R^2$ = 0.86) in the prediction of soil carbonates. The results were also corroborated from the founding obtained by Gomez et al. (2008) and Gomez et al. (2012), for which predictions of carbonate using PLS regression method have achieved values of $R^2$ of about 0.94 and 0.76 respectively, suggesting that both PLS and SVM can be consider as suitable prediction models for $CaCO_3$ prediction.

Summarily, predictions of total carbon ($TC_{CNS}$), organic carbon ($TOC_{WB}$, $OC_{CNS}$, LOI and HC) and inorganic carbon ($IC_{CNS}$ and $CaCO_3$) fraction contents for the studied soils appeared to be very performing in Vis-NIR spectral range. In particular, organic carbon fractions were best predicted by combining the second derivative preprocessing with the PLS regression model, especially for the hot water extracted carbon (HC) values. As shown in Figure 3.16 that depicts the scatterplots deriving from the best prediction models for validation datasets, HC parameter (Figure 3.16d) showed the least scatter of validated samples from the regression line,

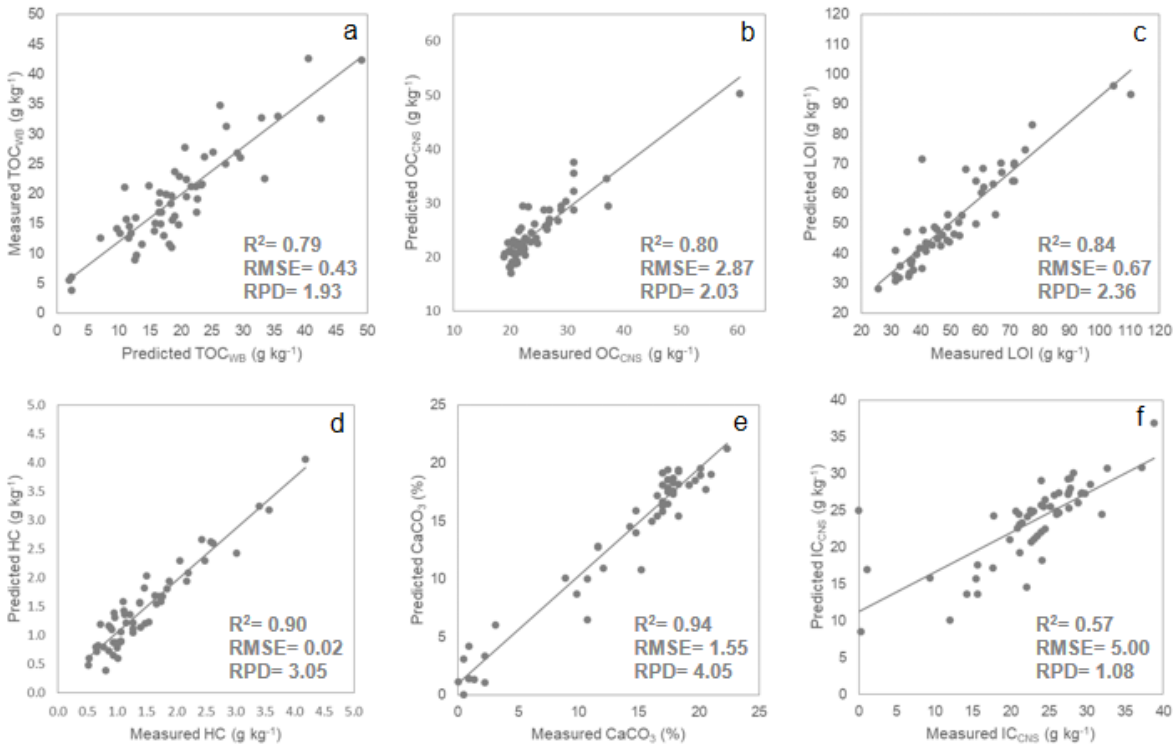


Figure 3.16 Scatter plots of Vis-NIR predicted versus measured values of soil carbon pools from validation dataset. Organic carbon fractions a) $TOC_{WB}$, b) $OC_{CNS}$, c) LOI and d) HC resulting from 2nd Der/PLS combination; inorganic carbon fractions e) $CaCO_3$ and f) $IC_{CNS}$ resulting from 1st Der/SVM combination.

as opposed to $TOC_{WB}$ (Figure 3.16a), which, among the organic carbon fractions, is the one that most exhibited a scattering effect and the lower values of prediction capability. At the same time, the combination that perfectly predicted inorganic carbon fractions was first derivative and SVM regression, particularly for $CaCO_3$ parameter, whose scatter plot (Fig. 3.16e) highlighted the great potential of Vis-NIR to predict both lower and higher values of soil carbonates. Instead, $IC_{CNS}$ parameter (Fig. 3.16f) showed the lowest values of $R^2$ and RPD and the highest RMSE in Vis-NIR prediction compared to the others carbon fractions, demonstrating the scatter plot that the estimation inaccuracy increased with decreasing $IC_{CNS}$ values.

From analysis of VIP (variable importance in the projection) profile, resulting from PLS regression model, it was possible to evaluate which wavelengths contributed most to the prediction of the various soil carbon fractions.

As resulting from Figure 3.17, which display the values of VIP>3 for soil parameters, the majority part of Vis-NIR spectra range had importantly contributed to predictions of all selected variables (soil carbon pools) in PLS regression model. The main absorbance regions selected in the PLS model were associated to soil organic and inorganic constituents in the Vis region at 400 nm (chromophores groups), in the NIR at 800 nm (organic pigments), 1400-1900 nm (OH groups), 2000-2200 nm (C-H and N-H groups), and 2200-2400 nm (C-H and C-O groups).

However, considerations about different contributions were made. First of all, the organic carbon fractions, in particular the hot water extracted organic carbon (HC), showed different relevant and continuous bands between 200 and 1200 nm, considered by Aïchi et al. (2009) as the most important for SOC prediction, and 1250-1500 nm that overlapped with C=O of carbonate. The secondary important spectral window for organic carbon of the studied soils,

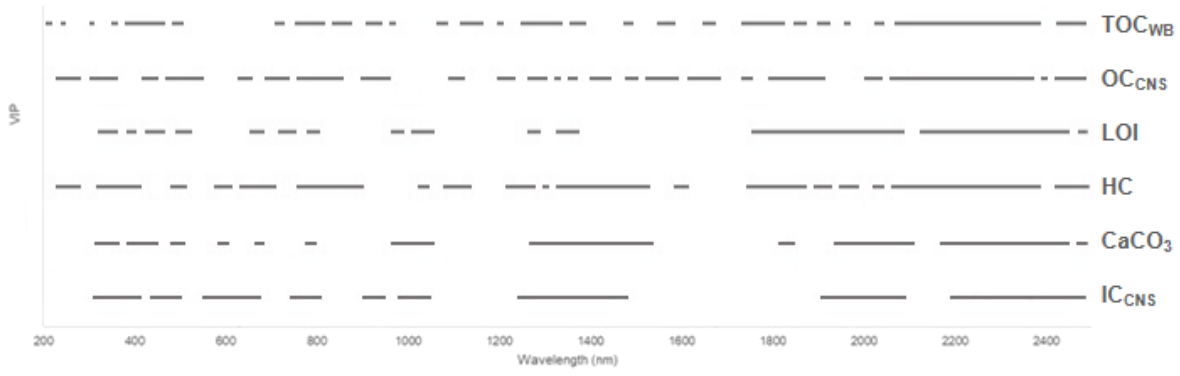


Figure 3.17 Summarizing of VIP profiles along the whole spectra range of the PLS regression model of organic and inorganic carbon fractions. Threshold of VIP>3 was taken into account to select the important wavelengths.

and this was particularly evident in LOI parameter, occurred in the two important part of the NIR region (1300-2100 nm and 2100-2500 nm). The last SWIR range considered by other authors as the most important for predicting SOC (Stenberg et al., 2010), was attributed to C-H and C-C bounds links in cellulose (Viscarra Rossel and Hicks, 2015). The absorption feature at 2200 nm is also due to the clay component because of lattice water, which is incorporated in the clay mineral structure. However, this absorption feature was also evident in the carbonate. Indeed, for soil inorganic carbon fractions (both $CaCO_3$ and $IC_{CNS}$) the VIP profile showed a sort of specular trend opposite to the organic fractions, although the Vis range also contributed well to inorganic fractions prediction. The SWIR region, from 1600 to 2500 nm, arranged the spectral bands that contributed most to their prediction. The VIP has achieved values between 5 and 6 for $CaCO_3$ and $IC_{CNS}$ fractions for the several bands around 2300-2500 nm, which contained signal of carbonates due to C-O stretching mode in $CaCO_3$ molecules (Ben-Dor and Banin, 1990; Riefolo et al., 2019; Asgari et al., 2020).

### 3.4 Validation of Vis-NIR predictions using external remote sensing Sentinel-2 data

With the aim to using the Vis-NIR regression model to extend the prediction of soil organic and inorganic carbon pools from field to farm-scale, satellite spectral data were used as external validation dataset of PLS regression models, previously obtained with laboratory soil spectra. External validation allowed predicting soil carbon pools in points where values were unknown. For the purpose, satellite Sentinel-2 spectra were chosen from multispectral imagery, downloaded by Copernicus Open Access Hub, for pixels of bare soil (NDVI threshold of 0.27).

Figure 3.18 shows raw reflectance spectra of laboratory soil spectra and interpolated row Sentinel-2 bare soil spectra (Figure 3.18a and b, respectively), for the 440-940 nm spectral

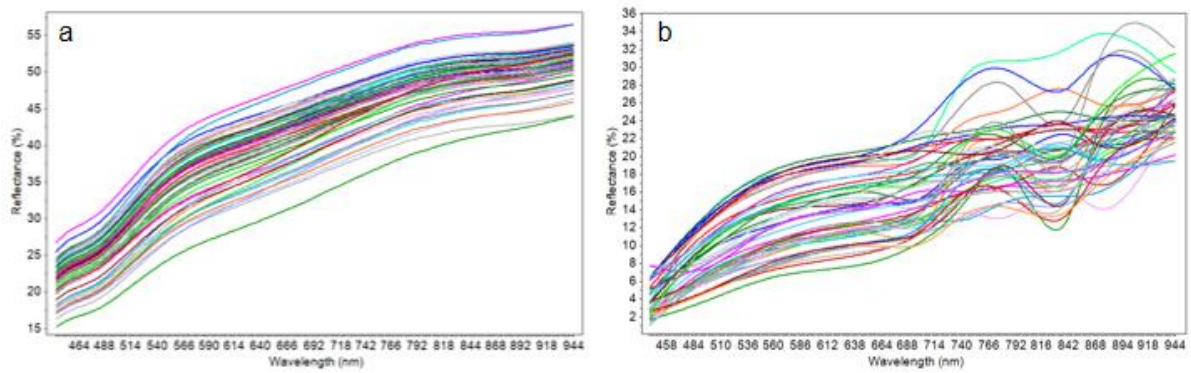


Figure 3.18 Raw reflectance spectra for the spectral range 440-940 nm of a) laboratory soil samples spectra and b) Sentinel-2 soil spectra of selected multispectral imagery pixels.

range. Prior to application of PLS regression model to predict unknown samples, the Sentinel-2 spectra were preprocessed using 2nd derivative transformation for estimation of organic C fractions and 1st derivative transformation for inorganic fractions, according to the preprocessing transformation that allowed achieving the best performances in PLS regression model, on which was discussed above.

Table 3.6 reports the percentage of the explained X total variance (wavelengths of Sentinel-2 spectra) resulting from prediction, according to seven different PLS factors. As noticeable, organic carbon fractions explained a lower variance of spectral dataset compared to the inorganic fractions. There was no high difference between the organic fractions, with an about 20% of explained total variance for $OC_{CNS}$ and LOI parameters, and 19% for $TOC_{WB}$ and HC.

The sensibility increased when inorganic carbon fractions were considered, with a percentage of about 52% in $CaCO_3$ content in correspondence of factor-6, and 56% for $IC_{CNS}$ parameter, according to factor-7. These findings suggested that predicted values of soil carbon pools using Sentinel-2 data were subject to a considerable estimation error, greater for the organic carbon values than for the inorganic carbon ones. The causes of the low estimation accuracy could be found in the differences or dissimilarities between the spectral signatures of the soils analyzed in the laboratory, compared to those downloaded from Sentinel-2, especially in the first part of NIR region (700-940 nm; Figure 3.18) (de Vries and Ter Braak Cajo, 1993).

The differences were mainly due to the higher heterogeneity in soil surface of Sentinel-2 samples, in terms of soil moisture and roughness, which most influenced the NIR bands of the soil spectra (Nanni and Dematte, 2006). These occurrences were completely avoided in laboratory spectra because of air-drying and grinding of soil samples prior to spectra acquisition, which standardized spectral measurements (Torrent and Barrón, 2008). Moreover,

Table 3.6
Values of explained X total variance in validation performances of Sentinel-2 spectra according to the factors of PLS regression model.

| | $TOC_{WB}$ | $OC_{CNS}$ | LOI | HC | $CaCO_3$ | $IC_{CNS}$ |
|---|---|---|---|---|---|---|
| Factor-1 | 17.6 | 17.7 | 17.1 | 17.2 | 12.5 | 13.5 |
| Factor-2 | 16.6 | 17.9 | 17.2 | 17.9 | 35.6 | 19.3 |
| Factor-3 | 17.4 | 17.7 | 18.8 | 18.2 | 40.1 | 39.0 |
| Factor-4 | 19.8 | 18.6 | 20.4 | 19.1 | 46.2 | 39.0 |
| Factor-5 | 19.3 | 17.9 | 20.1 | 18.9 | 51.0 | 45.2 |
| Factor-6 | 19.0 | 19.9 | 20.2 | 19.5 | 52.4 | 53.5 |
| Factor-7 | 19.6 | 19.6 | 19.2 | 18.8 | 46.9 | 56.3 |

atmospheric scattering and absorption along Vis-NIR range, which could cause lack of information (Hosseini Aria et al., 2012), mostly affect Sentinel-2 spectra. According to Shivangi and Kumari (2020), atmospheric correction becomes essential when working with remote sensing data, and the values of aerosol optical density increase during winter and summer seasons.

Most of the authors (Vaudour et al., 2019; Dvorakova et al., 2020; Piccini et al., 2020; Silvero et al., 2021; Zhou et al., 2021) that recently focused on soil carbon pools by remote sensing data, used Sentinel-2 soil spectra for both calibration and validation in the regression models, obtaining interesting results. However, the using of remote sensing data is particularly hampered when working in study areas where the principles of conservation agriculture (permanent soil cover) are applied, due to the lack of bare soil points. For this reason, in this case study it was preferred to conduct the analysis of the soils with the laboratory spectroscopy technique, and only at a later time to expand and integrate it through the remote sensing technique, to obtain a view at farm-scale and beyond of field.

Anyway, as possible to see from Figure 3.19 and 3.20 that summarize the correlations between organic and inorganic carbon fractions, the predicted values were highly correlated to

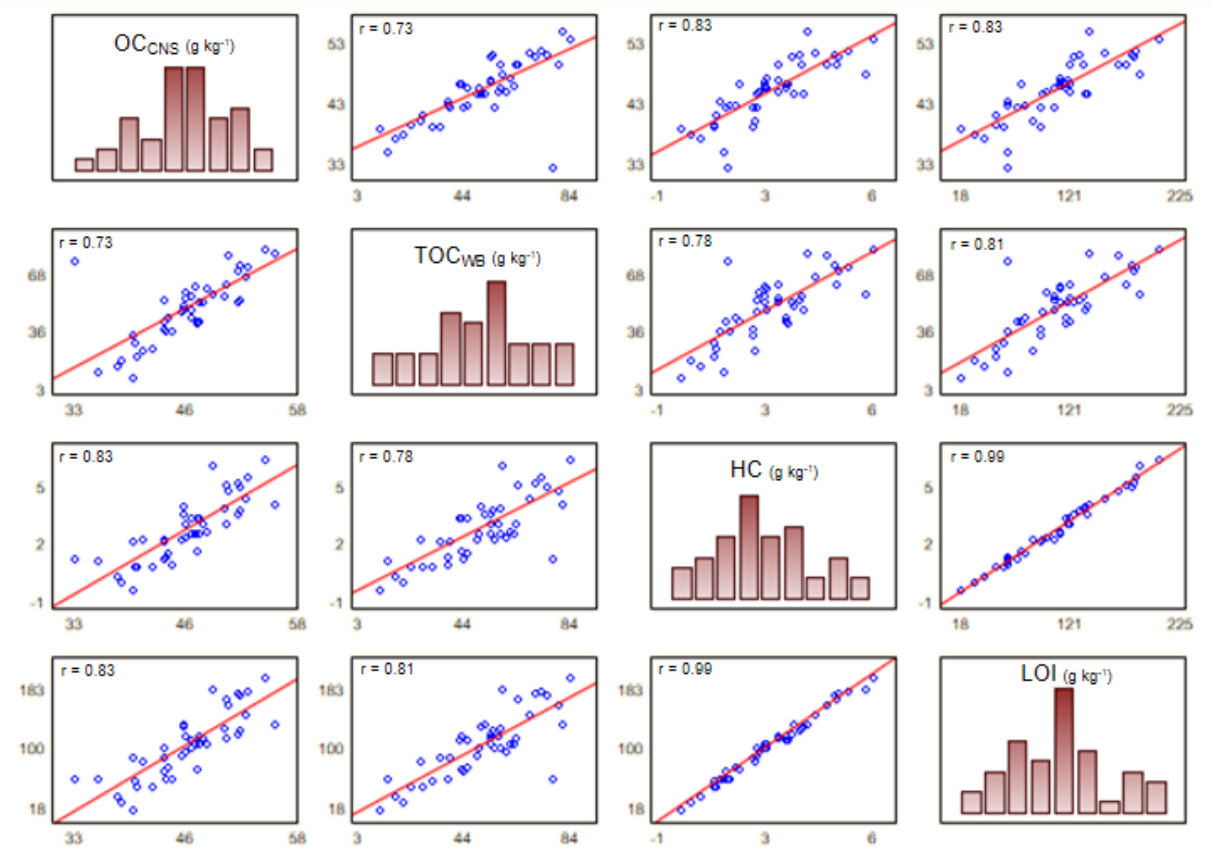


Figure 3.19 Pearson correlations (*p*-value <0.05) and frequency distribution between the Sentinel-2 predicted values of soil organic carbon fractions ($TOC_{WB}$, $OC_{CNS}$, LOI and HC).

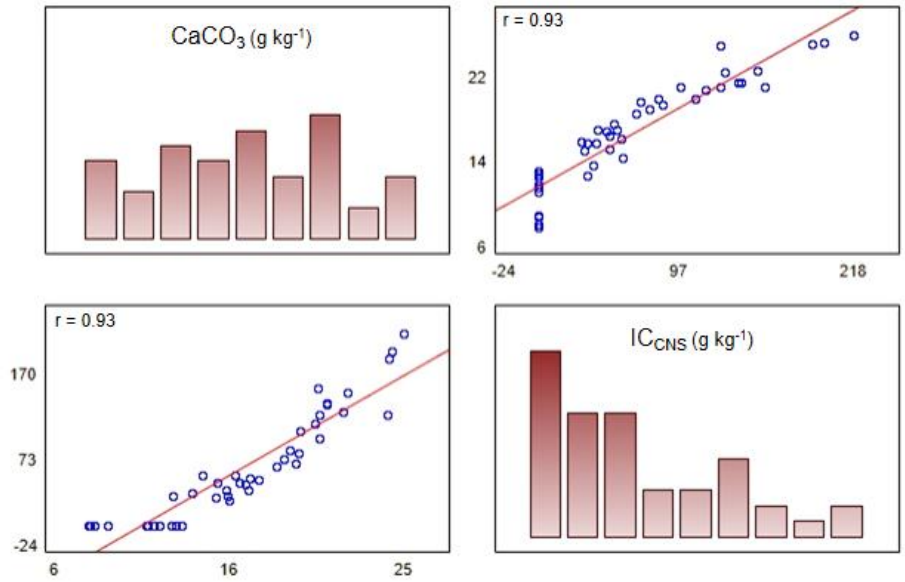


Figure 3.20 Correlations between the Sentinel-2 predicted values of soil inorganic carbon fractions ($CaCO_3$ and $IC_{CNS}$).

one each other, demonstrating that the Sentinel-2 has the capability to lower the gap between the various carbon fractions, found instead with the data measured thought laboratory methods. In particularly, predicted organic carbon fractions that showed all a normal distribution weakly skewed, exhibited a better correlation between $TOC_{WB}$ and $OC_{CNS}$, and $OC_{CNS}$ and LOI, with a linear regression coefficient very close to 1 in the scatter plot of LOI against HC parameter (Figure 3.19). The same also occurred for inorganic carbon fractions (Figure 3.20), although the predicted $IC_{CNS}$ parameter showed higher values compared to those of total carbonate (expressed in equivalents of C), which explained the partial less correspondence between the two for the lower values of inorganic carbon.

Sentinel-2 then proved to be a powerful tool capable of recognizing and differentiating the various soil carbon pools, especially those most exposed to the risk of degradation and erosion (HC).

### 3.4.1 Spatial prediction of Sentinel-2 carbon pools from field to farm-scale

Predicted Sentinel-2 soil carbon pools were subsequently used to create thematic maps of the entire territories of the farm analyzed, obtaining a view from field-scale to farm-scale with the sole contribution of integration between Vis-NIR spectroscopy and Sentinel2 multispectral data.

For spatial predictions of predicted soil carbon pools, geostatistical methods were used. In particular, according to previous results from spatialization of measured soil carbon pools,

exponential approach to geostatistics ordinary kriging was chosen. Figure 3.21 shows the spatial predictions of organic carbon pools of the soils, suggesting a large differentiation for the farm territories in terms of organic carbon contents. As could be seen, the fields that have been the subject of in-depth chemical investigation (in the central part of the map) showed the lowest values of organic carbon pools, with a high spatial correlation between them. Instead, the fields further south were those with the highest organic carbon values, probably due to the presence of wooded areas in the immediate vicinity.

The predicted inorganic carbon pools shown in Figure 3.22 exhibited a trend opposite to that of organic pools. In particular for carbonate content (Figure 3.22a), the fields of the northern part of the farm showed the lowest carbonate values, resulting completely decarbonated and

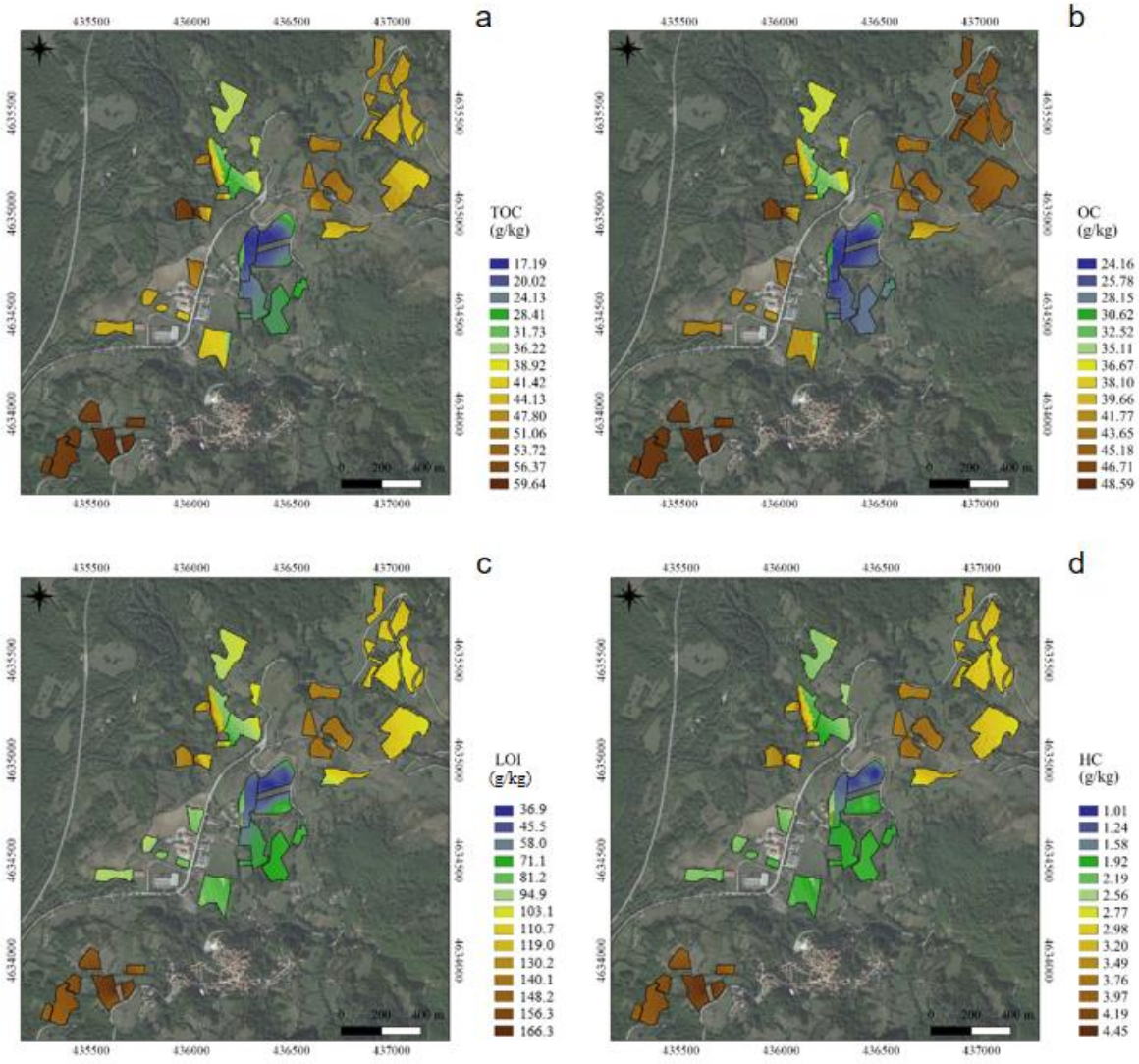


Figure 3.21 Spatial predictive soil maps with ordinary kriging of predicted Sentinel-2 organic carbon pools: a) $TOC_{WB}$, b) $OC_{CNS}$, c) LOI and d) HC across the Melise farm territories.

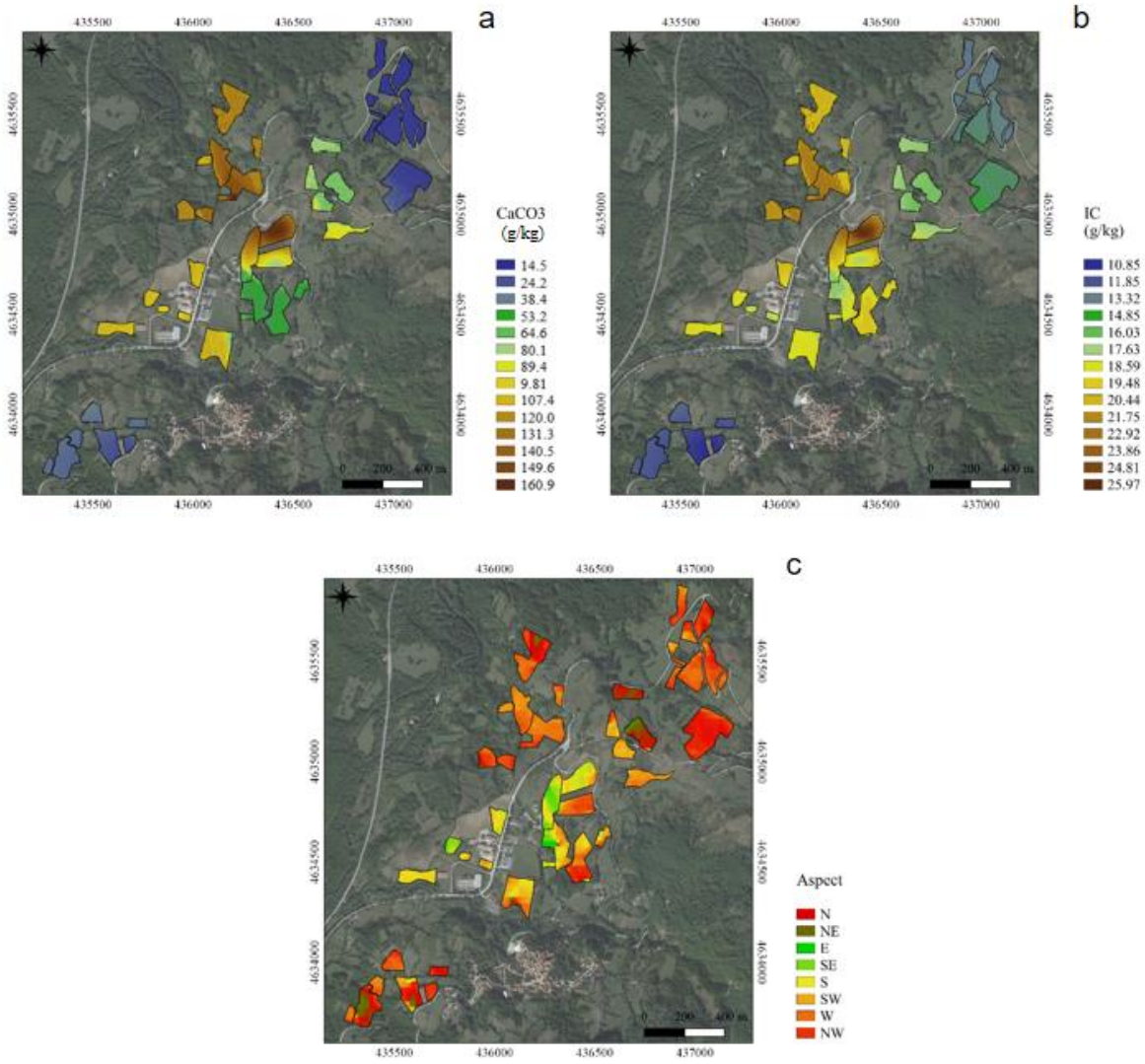


Figure 3.22 Spatial predictive soil maps of predicted Sentinel-2 inorganic carbon pools: a) $CaCO_3$ and b) $IC_{CNS}$. c) Aspect of Melise farm territories.

therefore subjected to a very fast carbonate dynamics. On the other hand, the fields subjected to chemical analysis (central part of the Melise territories) resulted in those with the highest inorganic carbon content (both $CaCO_3$ and $IC_{CNS}$ parameters). In the specific case of the area under study, the NDVI values, although indicating a good vegetative vigor (see Figure 3.9c), do not correspond to the productivity of the apple orchard, indicating that other physiological and climatic factors (failure to fruit set and late frosts) had a more marked impact than the carbonate content, reducing the yield.

Generally, even with the predicted values, it was possible to state that as carbonates increased, the organic content of soils decreased, especially in their most extreme values. However, soils that contained average values of calcium carbonate and inorganic carbon also had a good organic carbon content, demonstrating that carbonate dynamics was certainly not

the only factor that correlates with SOC content. Indeed, as clearly evident from Figure 3.22c, the factor that most influenced the dynamics of carbon was that of exposure, confirming that the fields with north exposure were those with the greatest contents of organic carbon (the lower in carbonate), and in which the dynamics of SOC was slower than elsewhere. Furthermore, to corroborate this result, the farm's soils that showed the highest values of organic carbon fractions were those that, prior to conversion into apple orchards, were made up of permanent meadows, and therefore already rich in SOC. Thus, the soil management factor appears to be predominant for the accumulation or loss of the organic carbon.

In the final analysis, by means of the semivariogram models obtained from the application of the exponential OK models, it was possible to evaluate the goodness of the spatial predictions of the values predicted by Sentinel-2. All soil carbon parameters reached spatial incorrelation values around 1200 m, with the exception of the total carbonate parameter (Figure 3.23e), which showed a long-range spatial autocorrelation (1600 m). The values of LOI and HC parameters (Figure 3.23c and d, respectively) showed the lower nugget effect, compared to those of $TOC_{WB}$ and $OC_{CNS}$ (Figure 3.23a and b, respectively). In general, for all soil carbon values, both organic and inorganic, the actual values (points in the semivariograms) predicted by Sentinel-2 showed a greater scattering effect than the fitted ones (solid line), compared with those measured for

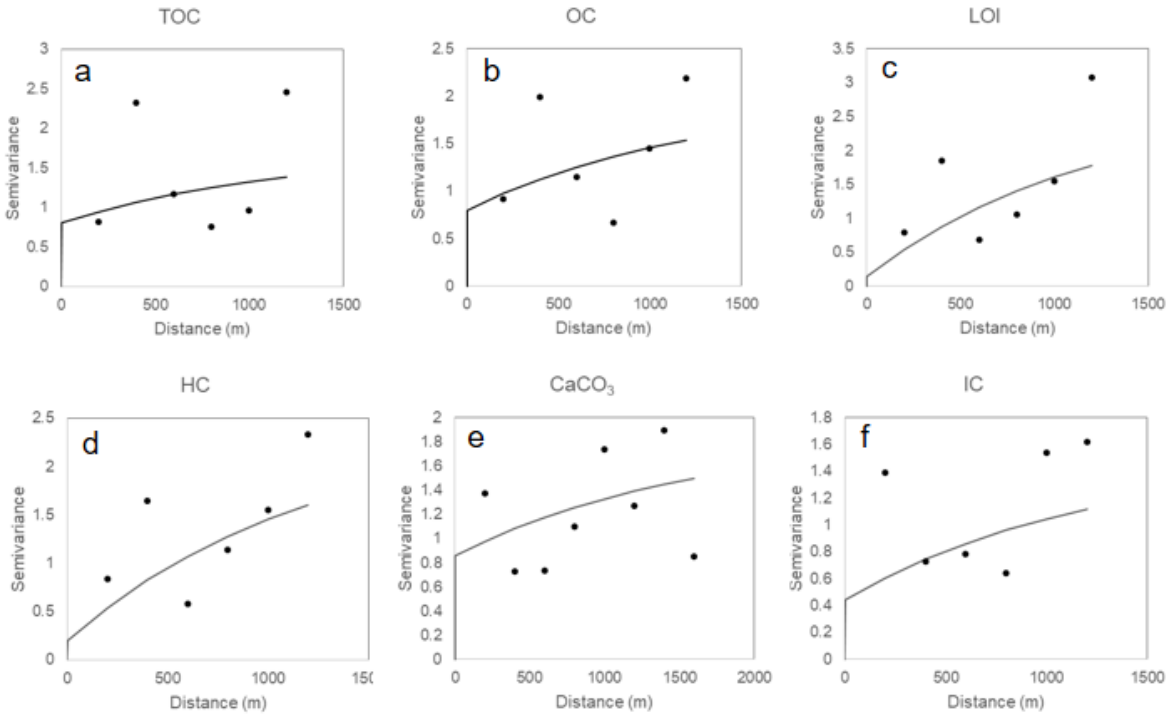


Figure 3.23 Experimental (points) and fitted (solid lines) model semivariograms of Sentinel-2 predicted values of a) $TOC_{WB}$, b) $OC_{CNS}$ c) LOI, d) HC, e) $CaCO_3$ and f) $IC_{CNS}$ parameters.

the study area, and therefore a greater spatial prediction error. This condition was mainly due to the different type of sampling, regular in the fields subject to chemical analysis, clusters type for the Sentinel-2 points, whose position was linked to the presence of bare soil. Indeed, as suggested by (Castrignanò et al., 2000), regular sampling leads to a better performance of the geostatistical models compared to a random and cluster sampling.

However, it was possible to observe that from the integration between Vis-NIR and Sentinel-2 techniques good prediction maps of the soil carbon pools were obtained, suggesting that the approach adopted in this research is valid for carrying out a screening of the soils subject to study and estimate their ability to lose or acquire organic carbon.

## Conclusions

From different carbon chemical analysis methods, it was possible to measure the different fraction of the soil organic and inorganic carbon from fields of an organic farm, in which the principles of conservation agriculture (permanent soil cover) are applied.

By the employing of CNS analyzer, total C and organic C content were assessed, namely respectively $TC_{CNS}$ and $OC_{CNS}$. The difference between them was indicated as total inorganic C ($IC_{CNS}$). From wet oxidation procedure turned out total organic C ($TOC_{WB}$) and hot water-extracted organic C (HC), the latter previously treated with hot water (80°C) and considered as soluble organic C fraction, as more labile and active in the soil. Moreover, loss on ignition (LOI) as dry combustion method was employed to deepen the quality of organic materials in the studied soils, and total calcium carbonate content by HCl treatment, as part of inorganic C fraction. Our result indicated that $OC_{CNS}$ and $TOC_{WB}$ parameters (the means values were 24.6 and 19.6 g $kg^{-1}$, respectively) accounted for 52% and 41% of total C ($TC_{CNS}$; mean of 47.5 g $kg^{-1}$). This suggested that $OC_{CNS}$ estimated an organic fraction of 20% greater than TOC, which was resistant (namely recalcitrant) to wet chemical oxidation. The largest part of SOC seemed to be contained in LOI parameter (mean of 51.4 g $kg^{-1}$), although LOI is well known to be overestimates, due to losses of structurally bound water in clay minerals and other non-SOC fractions. Moreover, $OC_{CNS}$ and $TOC_{WB}$ represented about 47% and 38% of LOI, indicating that the conventional LOI-to-SOC conversion factor of 0.58 was not applicable in the studied soils. As inorganic fractions, $IC_{CNS}$ and $IC_{CARB}$ (means values of 22.8 g $kg^{-1}$ and 15.2 g C-$CaCO_3$ $kg^{-1}$, respectively) were about 48% and 32% of $TC_{CNS}$, suggesting that $IC_{CNS}$ parameter contained 33% more inorganic carbon than $IC_{CARB}$, and which was different from calcium carbonate. The value of HC (mean of 1.5 g $kg^{-1}$), which constituted the soluble organic carbon fraction (active and labile), was positively correlated to both $TOC_{WB}$ and $OC_{CNS}$ values, and to biological properties. In particular, the increasing of labile pool of organic carbon resulted in an increase of microbial activity ($qCO_2$ index). The value of organic C accumulated (Stock C kg $m^{-2}$, mean of 4.45 stock C kg $m^{-2}$) seemed to increase the formation of macro-pores and macro-aggregates induced by the presence and management of grass cover. This finding indicated that long-term orchard-cropping farming system increased SOC stocks in the topsoil layer.

From a spatial prediction of measured soil C pools, ordinary kriging linked to the analysis of semivariograms, it was possible to split the studied area in two sub-area: to the north, with the lower values of organic C, and to the south, characterized by high organic C values. This appeared to be related to their different exposure and slope, with the higher organic C values for northern exposures of soils, showing slower carbon dynamics and greater carbon accumulation. HC contents increased with the increasing of organic C, but the $HC/TOC_{WB}$ ratio decreased at the increasing of organic C, justifying the faster carbon dynamics in the northern field. Inorganic C fractions showed an opposite trend compared to organic ones, classifying the soils from calcareous to weakly calcareous or completely decarbonated, due to the great variability of climate (increasing soil moisture more decarbonated) and more carbonate in the parent material. The biological properties (soil respiration and microbial activity) were found to be negatively affected by the presence of carbonates.

The application of Vis-NIR spectroscopy as proximal soil sensing technique allowed to obtain accurately results in predictive regression models for the studied soil properties. First, from a qualitative check of soil spectra it turned out that several large bands of absorbance in the SWIR, between 1400 and 2200 nm, were lied to the presence of many organic functional groups such as phenolic O-H, amide N-H, amine H and aliphatic C-H. In addition, spectra of hot water extracts revealed the typical characteristics of the organic functional groups, whose absorbance peaks has not been overlapped from absorptions of the inorganic carbon and mineral fractions, which were present in lower concentrations than in untreated soils. Inorganic C forms exhibited typical absorptions at 2245 and 2345 nm, due to C-O stretching mode in $CaCO_3$ molecules.

For the construction of the predictive models for the C pools parameters of the studied soils, three different regression methods were chosen, partial least squares (PLS), support vector machine (SVM) and boosted trees (BT). Previously to their application, the spectra were treated according to five preprocessing methods, Savitzky-Golay filter and absorbance transformation, first and second derivatives, standard normal variate (SNV) and multiplicative scatter correction (MSC) transformations. The organic C fractions had gained the best results from a combination between second derivative and PLS regression in validation dataset, with a performances of $R^2$=0.79, RMSE=0.43 and RPD=1.93 for $TOC_{WB}$, $R^2$=0.80, RMSE=2.87 and RPD=2.03 for $OC_{CNS}$, $R^2$=0.90, RMSE=0.02 and RPD=3.05 for HC, and $R^2$=0.68, RMSE=0.96 and RPD=1.51 for LOI. The inorganic C fractions ($IC_{CNS}$ and $CaCO_3$) as well as total C ($TC_{CNS}$), were optimally predicted in validation by the combination of second derivative transformation and SVM regression, with performances of $R^2$=0.58, RMSE=4.90 and

RPD=1.11 for $IC_{CNS}$, $R^2$=0.89, RMSE=2.17 and RPD=2.52 for $CaCO_3$, and $R^2$=0.83, RMSE=0.31 and RPD=1.90 for $TC_{CNS}$. Analysis of variable important in projection (VIP) allowed evaluating which wavelengths contributed most to the prediction of the various soil C fractions, suggesting a considerable effort from Vis spectral range for prediction of organic C forms. With an opposite trend, SWIR spectral range was the most important in prediction of inorganic C forms, for which the VIP has achieved values between 5 and 6 for $CaCO_3$ and $IC_{CNS}$ fractions for the several bands around 2300-2500 nm.

Lastly, in order to assess the soil C pools from field to farm scale, the integration with Vis-NIR spectroscopy and satellite Sentinel-2 data was proposed as validation in predictive models. Sentinel-2 spectra of bare soil, downloaded by Copernicus Open Access Hub, were used as external validation in previous Vis-NIR calibration models for the prediction of soil C pools in points where values were unknown. Results of validations suggested that predictions of soil organic and inorganic C fractions have been subjected to a considerable estimation error in terms of percentage of the explained X total variance (Sentinel-2 spectra wavelengths), with a values of about 20% for organic C and 50% for inorganic C, indicated an increased sensibility when inorganic C fractions were considered. Causes of this low estimation accuracy could be found in the differences between the bench Vis-NIR spectral signatures and those downloaded from Sentinel-2, continuously affected by heterogeneity in soil surface, as moisture and roughness. On the other and, application of remote sensing data is particularly hampered when working in study areas where the principles of conservation agriculture (permanent soil cover) are applied, due to the lack of bare soil points. On the contrary, using of soil laboratory analysis coupled with Vis-NIR spectroscopy and satellite data could be very useful with a precision agriculture approach. Anyway, the predicted soil C pools were highly correlated to one each other, proving that Sentinel-2 could be a powerful tool capable of recognizing and differentiating the various soil C pools, especially those most exposed to the risk of degradation and erosion.

# Appendix A

## Statistics details of measured soil carbon pools

The histogram plots of the measured organic and inorganic carbon fractions for the study areas are shown below. The main statistics are displayed.

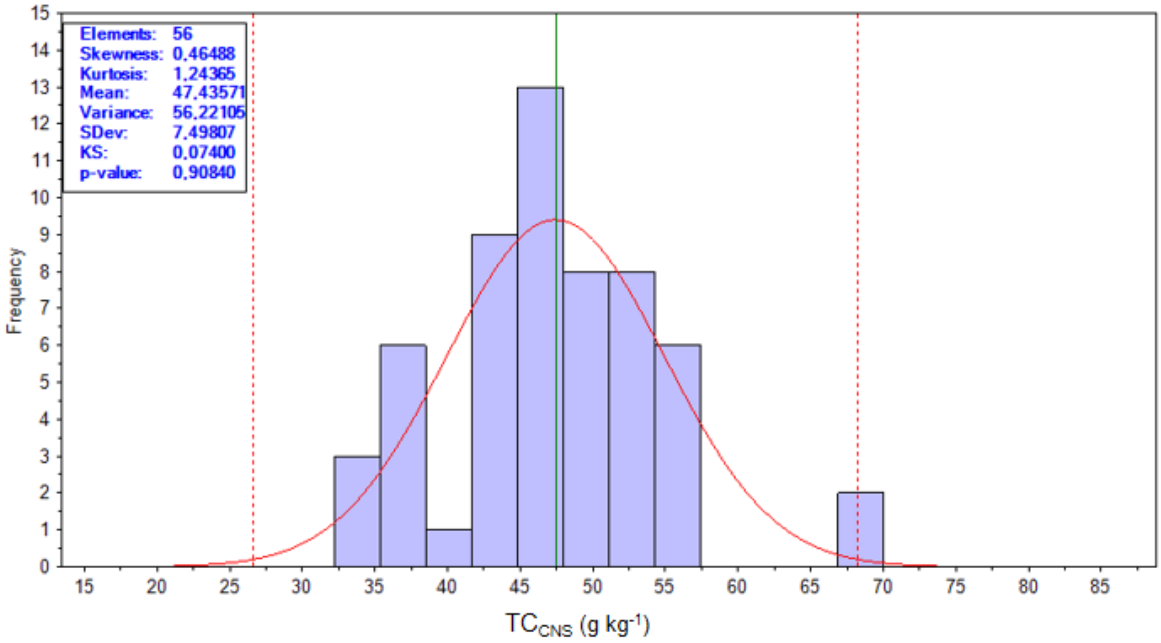


Figure A.1 Histogram plot of total C ($TC_{CNS}$) soil parameter, by CNS analyzer.

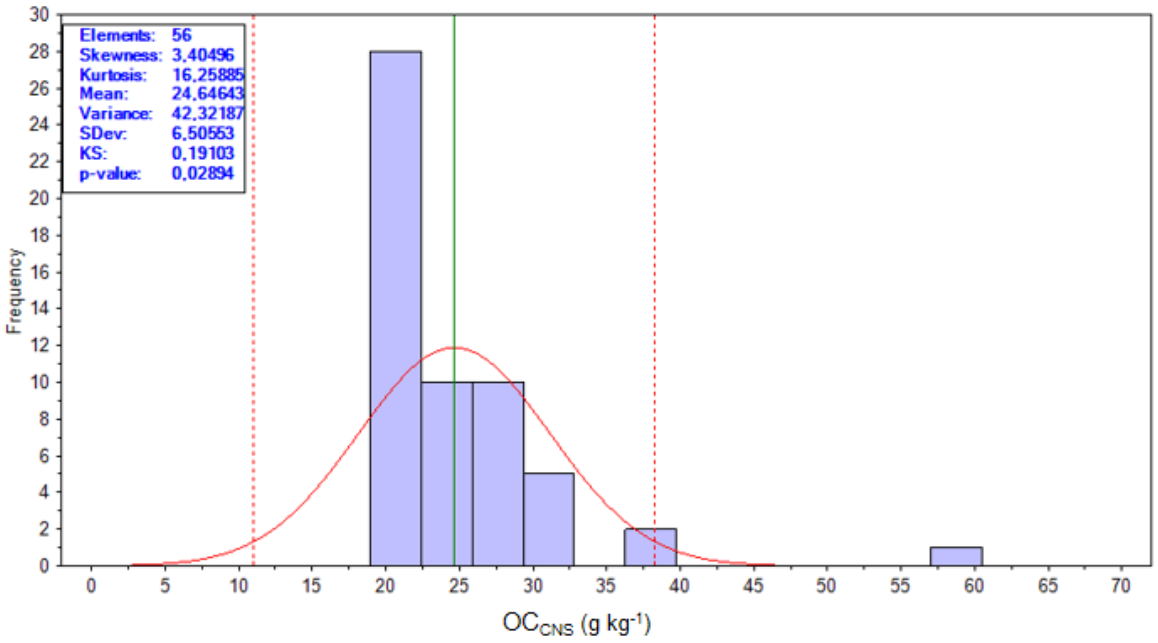


Figure A.2 Histogram plot of organic C ($OC_{CNS}$) soil parameter, by CNS analyzer.

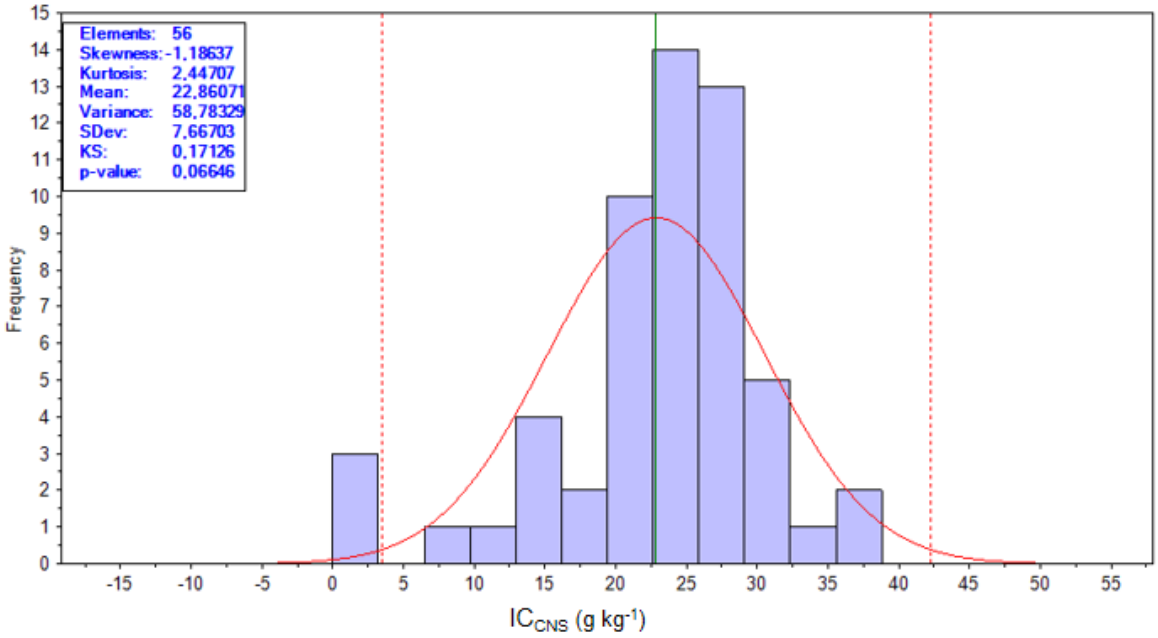


Figure A.3 Histogram plot of inorganic C ($IC_{CNS}$) soil parameter, as $TC_{CNS} – OC_{CNS}$ by CNS analyzer.

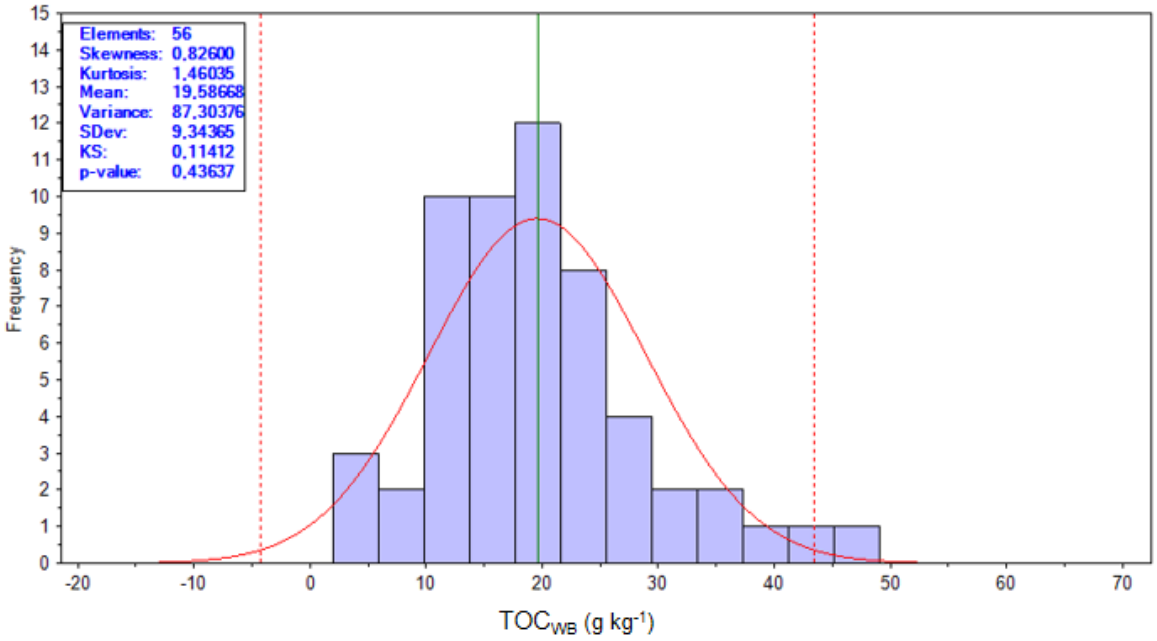


Figure A.4 Histogram plot of total organic C ($TOC_{WB}$) soil parameter, by wet oxidation method.

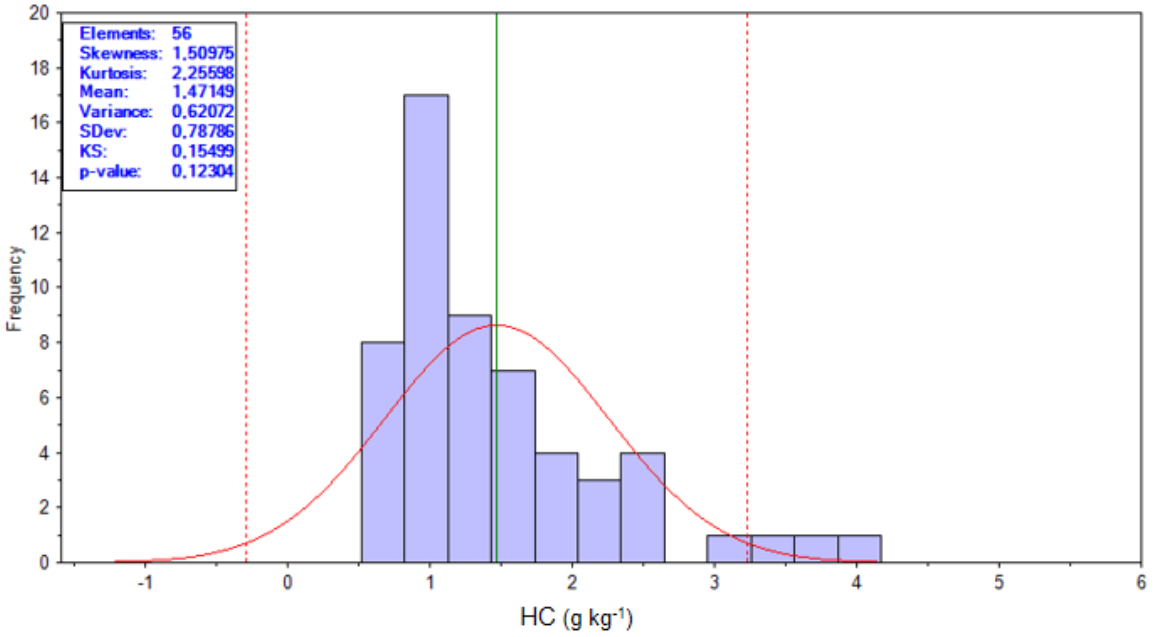


Figure A.5 Histogram plot of water-extracted organic C (HC) soil parameter, by wet oxidation method.

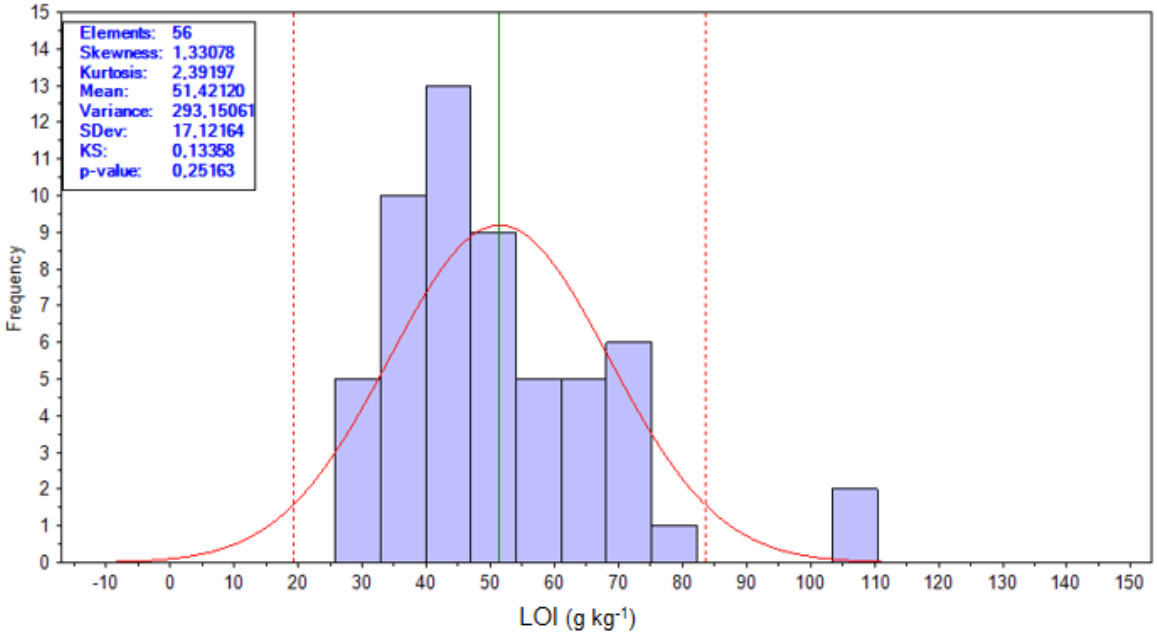


Figure A.6 Histogram plot of loss on ignition (LOI) soil parameter, by dry combustion method.

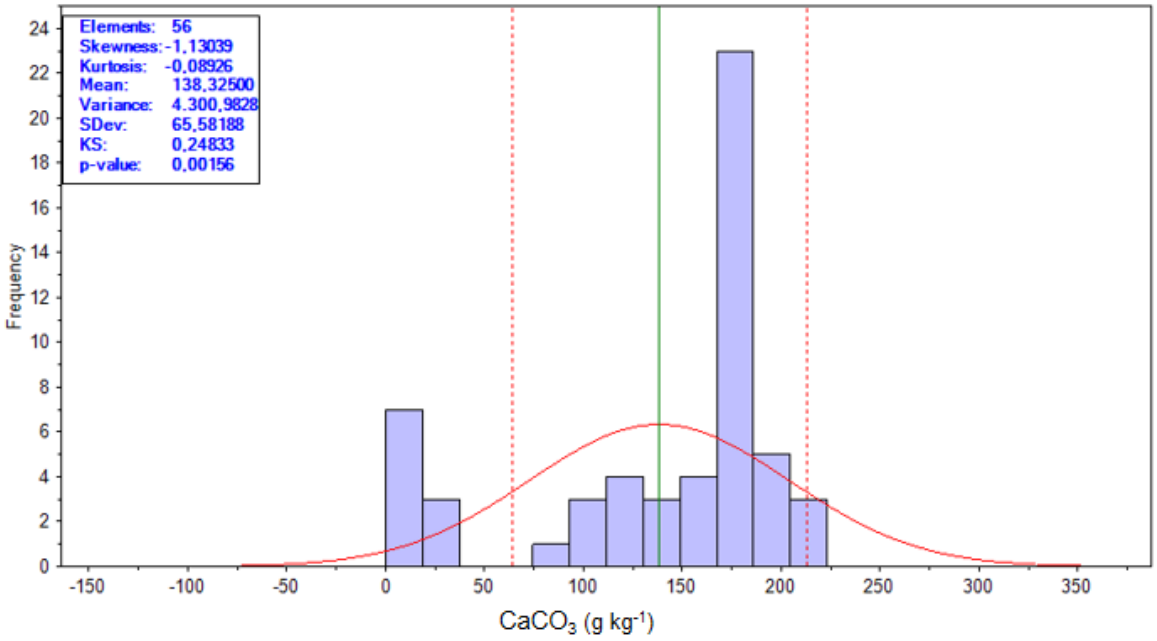


Figure A.7 Histogram plot of total calcium carbonate content ($CaCO_3$) soil parameter, by HCl treatment.

## List of References


Aceves M.B., Grace C., Hart M., Lin Q., Brookes P.C., 1994. Laboratory Manual of the Soil Microbial Biomass Group. Rothamsted Experimental Station, Soil Science Department, Harpenden, pp. 8–9.

Ahn M.Y., Zimmerman A.R., Comerford N.B., Sickman J.O., Grunwald S., 2009. Carbon Mineralization and Labile Organic Carbon Pools in the Sandy Soils of a North Florida Watershed. Ecosystems, 12: 672–685.

Aïchi H., Fouad Y., Walter C., Viscarra Rossel R.A., Lili Chabaane Z., Sanaa M., 2009. Regional predictions of soil organic carbon content from spectral reflectance measurements. Biosyst Eng., 104(3):442–446.

Aksoy E., Panagos P., Montanarella L., 2012. Spatial prediction of soil organic carbon of Crete by using geostatistics. Digital Soil Assessments and Beyond – Minasny, Malone & McBratney (eds), Taylor & Francis Group, London, ISBN 978-0-415-62155-7.

Álvaro-Fuentes J., López M.V., Cantero-Martinez C., Arrúe J.L., 2008. Tillage Effects on Soil Organic Carbon Fractions in Mediterranean Dryland Agroecosystems. Soil Sci. Soc. Am. J., 72:541–547.

Anderson T.H., Domsch K.H., 1993. The metabolic quotient for $CO_2$ ($qCO_2$) as a specific activity parameter to assess the effects of environmental conditions, such as pH, on the microbial biomass of forest soils. Soil Biol. Biochem., 25(3):393-395.

Angelopoulou T., Tziolas N., Balafoutis A., Zalidis G., Bochtis D., 2019. Remote Sensing Techniques for Soil Organic Carbon Estimation: A Review. Remote Sens.,11(6):676.

Asgari N., Ayoubi S., Demattê J.A.M., Dotto A.C., 2020. Carbonates and organic matter in soils characterized by reflected energy from 350-25000 nm wavelength. Journal of Mountain Science, 17(7).

Aucelli P.P.C., Robustelli G., Rosskopf C.M., Scarciglia F., Di Paola G., Lucà F., 2010. Geomorphological Map of the area between Frosolone and Trivento (Molise, Italy). Journal of Maps, 6:1, 423-434.

Barrón V., Torrent J., 1986. Use of the Kubelka-Munk theory to study the influence of iron oxides on soil colour. Journal of Soil Science, 37:499-510.

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

Baumgardner M.F., Silva L.R.F., Biehl L.L., Stoner E.R., 1986. Reflectance properties of soils. Adv Agron., 38:1-44.

Ben-Dor E., Banin A., 1989. Determination of organic matter content in arid-zone soils using a simple "loss-on-ignition" method. Communications in Soil Science and Plant Analysis, 20:15-16, 1675-1695.

Ben-Dor E., Banin A., 1990. Near-Infrared Reflectance Analysis of Carbonate Concentration in Soils. Applied Spectroscopy, 44(6):1064-1069.

Ben-Dor E., Banin A., 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Sci. Soc. Am. J., 59:364–372.

Ben-Dor E., Irons J.R., Epema G.F., 1999. Soil reflectance. In ''Remote Sensing for the Earth Sciences: Manual of Remote Sensing'' (A. N. Rencz, Ed.), Vol. 3, pp. 111–188. (3rd edn.). Wiley, New York.

Bernoux M., Feller C., Cerri C.C., Eschenbrenner V., Cerri C.E.P., 2006. Soil carbon sequestration. In: Roose E., Lal R., Feller C., Barthés B., Stewart B. (Eds.), Erosion and Carbon Dynamics. CRC Publisher, USA, pp. 13–22.

Bian R., Zhang Z., Zhang A., Zheng J., Li L., Joseph S., Pan G., Chang A., Zheng J., 2013. Effect of municipal biowaste biochar on greenhouse gas emissions and metal bioaccumulation in a slightly acidic clay rice paddy. BioResources, 9:685–703.

Bot A., Benites J., 2005. The Importance of Soil Organic Matter. FAO: Rome, Italy, p. 94.

Brown D.J., Bricklemyer R.S., Miller P.R., 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of vis–NIR soil C prediction in Montana. Geoderma, 129:251–267.

Buddenbaum H., Steffens M., 2012. The Effects of Spectral Pretreatments on Chemometric Analyses of Soil Profiles Using Laboratory Imaging Spectroscopy. Applied and Environmental Soil Science, 2012:274903, 1-12.

Cambule A.H., Rossiter D.G., Stoorvogel J.J., Smaling E.M.A., 2012. Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique. Geoderma, 183-184:41–48.

Castaldi F., Casa R., Castrignanò A., Pascucci S., Palombo A., Pignatti S., 2014. Estimation of soil properties at the field scale from satellite data: A comparison between spatial and non-spatial techniques. Eur. J. Soil Sci., 65:842–851.

Castrignanò A., Giugliarini L., Risaliti R., Martinelli N., 2000. Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate geostatistics. Geoderma, 97:39-60.

Chabbi A., Rumpel C., 2009. Guest editors' introduction. Organic matter dynamics in agro-ecosystems - the knowledge gaps. Eur. J. Soil Sci., 60:153–157.

Chang, C., Laird, D.A.2002. Near-infrared reflectance spectroscopic analysis of soil C and N. Soil Sci. 167:110–116.

Chang C.W., Laird D.A., Hurburgh Jr C.R., 2005. Influence of soil moisture on near-infrared reflectance Spectroscopic measurement of soil properties. Soil Science, 170(4):244-255.

Chang C.W., Laird D.A., Mausbach M.J., Hurburgh Jr. C.R., 2001. Near-infrared reflectance spectroscopy-principal components regression analysis of soil properties. Soil Science Society of America Journal, 65:480–490.

Chatterjee A., Lal R., Wielopolski L., Martin M.Z., Ebinger M.H., 2009. Evaluation of Different Soil Carbon Determination Methods, Critical Reviews in Plant Sciences, 28:3, 164-178.

Chilès J.P., Delfiner P., 2012. Geostatistics: modeling spatial uncertainty. 2nd ed. Hoboken (NJ): Wiley. Wiley Series in Probability and Statistics.

Christensen B.T., 1992. Physical fractionation of soil and organic matter in primary particle size and density separates. Advances in Agronomy, 20:1–90.

Christensen B.T., 2001. Physical fractionation of soil and structural and functional complexity in organic matter turnover. Eur. J. Soil Sci., 52:345–353.

Christy C.D., 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. Comput. Electron. Agric., 61(1):10-19.

Clark R.N., 1999. Spectroscopy of rocks and minerals and principles of spectroscopy. In ''Remote Sensing for the Earth Sciences'' (A. N. Rencz, Ed.), pp. 3–58. John Wiley & Sons, Chichester, UK.

Coates J., 2000. Interpretation of infrared spectra, a practical approach. In: "Encyclopedia of Analytical Chemistry", MEYERS R.A., Ed. Wiley & Sons Ltd.: Chichester, UK: 10815-10837.

Colombo C., Miano T., 2015. Metodi di analisi chimica del suolo. Ed. Pubblicità & Stampa, Bari (Italy).

Colombo C., Palumbo G., Di Iorio E., Sellitto V.M., Comolli R., Stellacci A.M., Castrignanò A., 2014. Soil organic carbon variation in alpine landscape (Northern Italy) as evaluated by diffuse reflectance spectroscopy. Soil Science Society of America Journal, 78:794–804.

Cozzolino D., Moron A., 2003. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. J. Agric. Sci. 140:65–71.

Danielson R.E., Sutherland P.L., 1986. Porosity. Methods of Soil Analysis, Physical and Mineralogical Methods. AS A, Madison, Wisconsin, pp. 443–461.

Davari M., Karimi S.A., Bahrami H.A., Hossaini S.M.T., Fahmideh S., 2021. Simultaneous prediction of several soil properties related to engineering uses based on laboratory Vis-NIR reflectance spectroscopy. Catena, 197:104987.

De Deyn G.B., 2015. Ecosystem Carbon and Soil Biodiversity. In: Ecosystem Services and Carbon Sequestration in the Biosphere. Lal R., Lorenz K., Hüttl F.R., Schneider B.U., von Braun J. (Eds), Springer Dordrecht Heidelberg New York London.

de la Paz Jimenez M., de la Horra A., Pruzzo L., Palma M.R., 2002. Soil quality: a new index based on microbiological and biochemical parameters. Biol. Fert. Soils, 35(4):302–306.

de Menezes M.D., Godinho Siulva S.H., de Mello C.R., Owens P.R., Curi N., 2016. Spatial prediction of soil properties in two contrasting physiographic regions in Brazil. Sci. Agric., 73(3):274-285.

De Vos B., Letterns S., Muys B., Deckers J.A., 2007. Walkley-Black analysis of forest soil organic carbon: recovery, limitations and uncertainty. Soil Use Manage., 23:221–229.

De Vries S., Ter Braak Cajo J.F., 1993. Prediction error in partial least squares regression: a critique on the deviation used in The Unscrambler. Chemometrics and Intelligent Laboratory Systems, 30:239-245.

Degens B.P., Schipper L.A., Sparling G.P., Duncan L.C., 2001. Is the microbial community in a soil with reduced catabolic diversity less resistant to stress or disturbance? Soil Biol. Biochem., 33:1143–1153.

Dhillon G.S., Gillespie A., Peak D., Van Rees K.C.J., 2017. Spectroscopic investigation of soil organic matter composition for shelterbelt agroforestry systems. Geoderma, 298:1–13.

Di Iorio E., Circelli L., Lorenzetti R., Costantini E.A.C., Egendorf S.P., Colombo C., 2019. Estimation of andic properties from Vis-NIR diffuse reflectance spectroscopy for volcanic soil classification. Catena, 182:104109.

Ding X.L., Han X.Z., Liang Y., Qiao Y.F., Li L.J., Li N., 2012. Changes in soil organic carbon pools after 10 years of continuous manuring combined with chemical fertilizer in a Mollisol in China. Soil Till. Res., 122:36–41.

Doetterl S., Stevens A., Six J., Merckx R., Van Oost K., Casanova Pinto M., Casanova-Katny A., Muñoz C., Boudin M., Zagal Venegas E., Boeckx P., 2015. Soil carbon storage controlled by interactions between geochemistry and climate. Nature Geosci., 8:780–783.

Dotto A.C., Dalmolin R.S.D., ten Caten A., Grunwald S., 2018. A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. Geoderma, 314:262–274.

Du C., Zhou J., 2009. Evaluation of soil fertility using infrared spectroscopy: A review. Environ. Chem. Lett., 7:97-113.

Dunn B.W., Beecher H.G., Batten G.D., Ciavarella S., 2002. The potential of near-infrared reflectance spectroscopy for soil analysis-a case study from the Riverine Plain of south-eastern Australia. Aust. J. Exp. Agric. 42, 607–614.

Dvorakova K., Shi P., Limbourg Q., van Wesemael B., 2020. Soil Organic Carbon Mapping from Remote Sensing: The effect of Crop Residues. Remote Sensing, 12:1913.

FAO, 2017. Soil Organic Carbon: the hidden potential. Food and Agriculture Organization of the United Nations. Rome, Italy.

FAO and ITPS, 2015. Status of the World's Soil Resources (SWSR) – Main Report. Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy.

Fernandez R.N., Schulze D.G., 1987. Calculation of soil color from reflectance spectra. Soil Science Society of America Journal, 51:1277-1282.

Ferré C., Castrignanò A., Comolli R., 2018. Comparison between spatial and non-spatial regression models for investigating tree–soil relationships in a polycyclic tree plantation of Northern Italy and implications for management. Agroforest Syst., https://doi.org/10.1007/s10457-018-0334-3.

Fidêncio P.H., Poppi R.J., Andrade J.C., Cantarella H., 2002. Determination of organic matter in soil using near-infrared spectroscopy and partial least squares regression. Communications in Soil Science and Plant Analysis, 33:1607–1615.

Frazier B.E., Cheng Y., 1989. Remote sensing of soils in the Eastern Palouse region with landsat thematic mapper. Remote Sens. Environ., 28:317–325.

Friedman J., Hastie T., Tibshirani R., 2000. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Ann. Stat. 28(2):337–407.

Froment A., 1972. Soil respiration in a mixed oak forest. Oikos 23, 273–277.

García-Díaz A., Marqués M.J., Sastre B., Bienes R., 2018. Labile and stable soil organic carbon and physical improvements using groundcovers in vineyards from central Spain. Sci. Total Environ, 621:387–397.

Gao B., 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 58(3):257-266.

Gao Y., Cui L., Lei B., Zhai Y., Shi T., Wang J., Chen Y., He H., Wu G., 2014. Estimating Soil Organic Carbon Content with Visible-Near-Infrared (Vis-NIR) Spectroscopy. Applied Spectroscopy, 68(7):712-722.

Geladi P., MacDougall D., Martens H., 1985. Linearization and scatter-correction for near-infrared reflectance spectra of meat. Applied Spectroscopy, 39:491–500.

Gholizadeh A., Borůvka L., Saberioon M.M., Kozàk J., Vašàt R., Němeček K., 2015. Comparing different data preprocessing methods for monitoring soil heavy metals based on soil spectral features. Soil and Water Research, 10:218-227.

Gholizadeh A., Žižal, D., Saberioon M., Boru°vka L., 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. Remote Sens. Environ., 218:89–103.

Gomarasca M., 2016. Il Telerilevamento. GEOmedia, 2(2):4-9.

Gomez C., Lagacherie P., Coulouma G., 2008. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. Geoderma, 148(2):141-148.

Gomez C., Lagacherie P., Coulouma G., 2012. Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis-NIR data. Geoderma, 189-190:176-185.

Guo L.B., Gifford R.M., 2002. Soil carbon stocks and land use change. Global Change Biol., 8:345–360.

Haddaway N.R., Hedlund K., Jackson L.E., Kätterer T., Lugato E., Thomsen I.K., Jørgensen H.B., Isberg P., 2016. How does tillage intensity affect soil organic carbon? A systematic review protocol. Environmental Evidence, 5:1.

Hayes M.H.B., Clapp C.E., 2001. Humic substances: considerations of compositions, aspects of structure, and environmental influences. Soil Science, 166:723–737.

Haynes R.J., 2005. Labile organic matter fractions as central components of the quality of agricultural soils: an overview. Advances in Agronomy, 85:221–268.

Helfrich M., Flessa H., Mikutta R., Dreves A., Ludwig B., 2007. Comparison of chemical fractionation methods for isolating stable soil organic carbon pools. European Journal of Soil Science, 58:1316–1329.

Henderson T.L., Baumgardner M.F., Franzmeier D.P., Scott D.E., Coster D.C., 1992. High dimensional reflectance analysis of soil organic matter. Soil Sci Soc Am J., 56:865–872.

Hoogsteen M.J.J., Latinga E.A., Bakker E.J., Groot J.C.J., Tittonell P.A., 2015. Estimating soil organic carbon through loss on ignition: effects of ignition conditions and structural water loss. European Journal of Soil Science, doi: 10.1111/ejss.12224.

Hosseini Aria S.E., Gorte B., Menenti M., 2012. Evaluation of sentinel-2 bands over the spectrum. Proc. 'First Sentinel-2 Preparatory Symposium', Frascati, Italy 23–27 April 2012, ESA SP-707.

Igne B., Reeves III J.B., McCarty G., Hively W.D., Lund E., Hurburgh Jr C.R., 2010. Evaluation od spectral pretreatments, partial least squares, least squares support vector machine and locally weighted regression for quantitative spectroscopic analysis of soils. J. Near Infrared Spectrosc., 21:495–509.

IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.

Jafarian Z., Kavian A., 2013. Effects of Land-Use Change on Soil Organic Carbon and Nitrogen. Communications in Soil Science and Plant Analysis, 44(1-4):339-346.

Janik L.J., Merry R.H., Skjemstad J.O., 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? Aust. J. Exp. Agric., 38:681-696.

Janik L.J., Skjemstad J.O., Shepherd K.D., Spouncer L.R., 2007. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. Australian Journal of Soil Research, 45:73–81.

Jensen J.L., Christensen B.T., Schjønning P., Watts C.W., Munkholm L.J., 2018. Converting loss-on-ignition to organic carbon content in arable topsoil: pitfalls and proposed procedure. European Journal of Soil Science, 69(4):604.

Jia X., O'Connor D., Shi Z., Hou D., 2021. VIRS based detection in combination with machine learning for mapping soil pollution. Environmental Pollution, 268(1):115845.

Kalcsits L., Lotze E., Tagliavini M., Hannam K.D., Mimmo T., Neilsen D., Neilsen G., Atkinson D., Casagrande Biasuz E., Borruso L., Cesco S., Fallahi E., Pii Y., Valverdi N.A., 2020. Recent Achievements and New Research Opportunities for Optimizing Macronutrient Availability, Acquisition, and Distribution for Perennial Fruit Crops. Agronomy, 10:1738.

Kämpf I., Hölzel N., Störrle M., Broll G., Kiehl K., 2016. Potential of temperate agricultural soils for carbon sequestration: A meta-analysis of land-use effects. Sci. Total Environ., 566–567:428–435.

Kirschbaum M.U.F., 2000. Will changes in soil organic carbon act as a positive or negative feedback on global warming? Biogeochemistry, 48:21–51.

Kooistra L., Wehrens R., Leuven R.S.E.W., Buydens L.M.C., 2001. Possibilities of visible-near-infrared spectroscopy for the assessment of soil contamination in river floodplains. Analytica Chimica Acta, 446(1-2):97–105.

Kortüm G., 1969. Reflectance spectroscopy. Springer-Verlag, Berlin.

Krull E.S., Skjemstad J.O., Baldock J.A., 2004. Functions of Soil Organic Matter and the Effect on Soil Properties; Report Project No CSO 00029; CSIRO, Grains Research & Development Corporation: Glen Osmond, Australia, 2004.

Laekemariam F., Kibret K., Shiferaw K., 2018. Potassium (K)-to-magnesium (Mg) ratio, its spatial variability and implications to potential Mg-induced K defciency in Nitisols of Southern Ethiopia. Agriculture & Food Security, 7:13.

Lal R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science, 304:1623–1627.

Lal R., 2008. Soil carbon stocks under present and future climate with specific reference to European ecoregions. Nutr. Cycl. Agroecosyst., 81:113–127.

Lal R., 2015. Sequestering carbon and increasing productivity by conservation agriculture. J. Soil Water Conserv., 70:55A–62A.

Lal R., 2017. Soil organic carbon sequestration: Importance and state of science. Proceedings of the Global Symposium on Soil Organic Carbon, Food and Agriculture Organization of the United Nations. Rome, Italy.

Laslett G.M., McBratney A.B., Pahl B.J., Hutchinson M.F., 1987. Comparison of several spatial prediction methods for soil pH. Soil Science, 38: 325-341.

Lee H., Ghosh S.K., 2009. Performance of information criteria for spatial models. Journal of Statistical Computation and Simulation, 79:1, 93-106.

Lehmann J., Kleber M., 2015. The contentious nature of soil organic matter. Nature, 528:60–68.

Li J., Wen Y., Li X., Li Y., Yang X., Lin Z., Song Z., Cooper J.M., Zhao B., 2018. Soil labile organic carbon fractions and soil organic carbon stocks as affected by long-term organic and mineral fertilization regimes in the North China Plain. Soil and Tillage Research, 175:281-290.

Liakos K.G., Busato P., Moshou D., Pearson S., Bochtis D., 2018. Machine learning in agriculture: A review. Sensors, 18:2674.

Lima D.L., Santos S.M., Scherer H.W., Schneider R.J., Duarte A.C., Santos E.B., Esteves V.I., 2009. Effects of organic and inorganic amendments on soil organic matter properties. Geoderma, 150:38–45.

Lindsay W.L., Norvell W.A., 1978. Development of a DTPA soil test for zinc, iron, manganese, and copper. Soil Science Society of America Journal, 42:421-428.

Linsler D., Sawallisch A., Höper H., Schmidt H., Vohland M., Ludwig B., 2017. Near-infrared spectroscopy for determination of soil organic C, microbial biomass C and C and N fractions in a heterogeneous sample of German arable surface soils. Archives of agronomy and soil science, 63(11):1499–1509.

Liu C., Huang P. M., 2002. Role of hydroxyl-aluminosilicate ions (proto-imogolite soil) in the formation of humic substances. Organic Geochemistry, 33:295-305.

Liu C., Li Z., Berhe A.A., Xiao H., Liu L., Wang D., Peng H., Zeng G., 2019. Characterizing dissolved organic matter in eroded sediments from a loess hilly catchment using fluorescence EEM-PARAFAC and UV-Visible absorption: insights from source identification and carbon cycling. Geoderma, 334:37-48.

Loiseau T., Chen S., Mulder V.L., Román Dobarco M., Richer-de-Forges A.C., Lehmann S., Bourennane H., Saby N.P.A., Martin M.P., Vaudour E., et al., 2019. Satellite data integration for soil clay content modelling at a national scale. Int. J. Appl. Earth Obs. Geoinform. 82, 101905.

Lucà F., Conforti C., Castrignanò A., Matteucci G., Buttafuoco G., 2017. Effect of calibration set size on prediction at local scale of soil carbon by Vis-NIR spectroscopy. Geoderma, 288:175–183.

Luo Z., Wang E., Sun O.J., 2010. Can no-tillage stimulate carbon sequestration in agricultural soils? A meta-analysis of paired experiments. Agric. Ecosyst. Environ., 139(1-2):224–231.

Madari B.E., Reeves III J.B., Coelho M.R., Machado P.L.O.A., De-Polli H., Coelho R.M., Benites V.M., Souza L.F., McCarty G.W., 2005. Mid and Near-Infrared Spectroscopic Determination of Carbon in a Diverse Set of Soils from the Brazilian National Soil Collection. Spectroscopy Letters: An International Journal for Rapid Communication, 38:6,721-740.

Magdoff F., Van E.H., 2010. Building Soils for Better Crops. 3rd ed., SARE Outreach Publications: Brentwood, CA, USA.

Mammadov E., Denk M., Riedel F., Lewinska K, Kaźmierowski C., Glaesser C., 2020. Visible and Near-Infrared Reflectance Spectroscopy for Assessment of Soil Properties in the Caucasus Mountains, Azerbaijan, Communications in Soil Science and Plant Analysis, DOI: 10.1080/00103624.2020.1820027.

Marschner B., Kalbitz K., 2003. Controls of bioavailability and biodegradability of dissolved organic matter in soils. Geoderma, 113:211–235.

Massaccesi L., De Feudis M., Agnelli A.E., Nasini L., Regni L., D'Ascoli R., Castaldi S., Proietti P., Agnelli A., 2018. Organic carbon pools and storage in the soil of olive groves of different age. European Journal of Soil Science, 69(5).

McCarty G.W., Reeves III J.B., Reeves V.B., Follet R.F., Kimble J.M., 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Science of America Journal, 66:640–646.

McLauchlan K.K., Hobbie H.S., 2004. Comparison of Labile Soil Organic Matter Fractionation Techniques. Soil Sci. Soc. Am. J. 68:1616–1625.

Mikutta R., Kleber M., Torn M.S., Jahn R., 2006. Stabilization of soil organic matter: association with minerals or chemical recalcitrance? Biogeochemistry, 77:25–56.

Mikutta, R., Kaiser, K. 2011. Organic matter bound to mineral surfaces: Resistance to chemical and biological oxidation. Soil Biol. Biochem. 43, 1738-1741.

Miller C.E., 2001. Chemical principles of near-infrared technology. In ''Near-Infrared Technology in the Agricultural and Food Industries'' (P. Williams and K. Norris, Eds.), pp. 19–37. The American Association of Cereal Chemists Inc., St. Paul, MN.

Miloš B., Bensa A., 2018. Prediction of organic carbon and calcium carbonates in agricultural soils with Vis-NIR spectroscopy. Poljoprivreda, 24(1):45–51.

Minasny B., Sulaeman Y., McBratney A.B., 2011. Is soil carbon disappearing? The dynamics of soil organic carbon in Java. Global Change Biol., 17:1917–1924.

Ministero per le Politiche Agricole e Forestali, 1999. Metodi ufficiali di analisi chimica del suolo. (ed. P. Violante). Franco Angeli Editore, Roma. D.M. del 13/09/99, Gazzetta Ufficiale n. 248 del 21.10.99.

Næs T., Isakson T., Fearn T., Davies T., 2002. A user-Friendly guide to Multivariate Calibration and Classification. Chichester (UK): NIR Publications.

Nanni M.R., Dematte, M.J.A., 2006. Soil line behavior obtained by laboratorial spectroradiometry for different soil classes. Revista Brasileira De Ciencia Do Solo, 30:1031–1038.

Nawar S., Buddenbaum H., Hill J., Kozak J., Mouazen A.M., 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. Soil & Tillage Research, 155:510–522.

Navarro-Pedreño J., Almendro-Candel M.B., Zorpas A.A., 2021. The Increase of Soil Organic Matter Reduces Global Warming, Myth or Reality? Sci, 3(18).

Nelson D.W., Sommers L.E., 1996. Chapter 34. p 1001-1006. Total Carbon, Organic Carbon, and Organic Matter. In: J.M. Bigham et al. (ed.) Soil Science Society of America and

American Society of Agronomy. Methods of Soil Analysis. Part 3. Chemical Methods-SSSA Book Series no. 5. Madison, WI.

Nocita M., Stevens A., Noon C., van Wesemael B., 2013. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. Geoderma, 199:37-42.

Ogle S.M., Breidt F.J., Paustian K., 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. Biogeochemistry, 72:87–121.

Ontl T.A., Schulte L.A., 2012. Soil Carbon Storage. Nature Education Knowledge, 3(10):35.

Paul E.A., 2014. Soil microbiology, ecology and biochemistry. Academic press. 598 p.

Paul E.A., Follett R.F., Leavitt S.W., Halvorson A., Peterson G.A., Lyon D.J., 1997. Radiocarbon dating for determination of soil organic matter pool sizes and fluxes. Soil Sci.Soc. Amer.J., 61: 1058-1067.

Paustian K., Lehmann J., Ogle S., Reay D., Robertson G.P., Smith P., 2016. Climate-smart soils. Nature, 532: 49-57.

Pham H., 2019. A New Criterion for Model Selection. Mathematics 2019, 7, 1215.

Piccini C., Francaviglia R., Marchetti A., 2020. Predicted Maps for Soil Organic Matter Evaluation: The Case of Abruzzo Region (Italy). Land, 9:349.

Poeplau C., Don A., Dondini M., Leifeld J., Nemo R., Schumacher J., Senapati N., Wiesmeier M., 2013. Reproducibility of a soil organic carbon fractionation method to derive RothC carbon pools. Eur. J. Soil Sci., 64:735–746.

Poeplau C., Don A., Six J., Kaiser M., Benbi D., Chenu C., Cotrufo M. F., Derrien D., Gioacchini P., Grand S., Gregorich E., Griepentrog M., Gunina A., Haddix M., Kuzyakov Y., Kühnel A., Macdonald L.M., Soong J., Trigalet S., Vermeire M.L., Rovira P., van Wesemael B., Wiesmeier M., Yeasmin S., Yevdokimov I., Nieder R., 2018. Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison. Soil Biology and Biochemistry, 125:10–26.

Post W.M., Izaurralde R.C., Jastrow J.D., McCarl B.A., Amonette J.E., Bailey V.L., Jardine P.M., Tristram O.W., Zhou J., 2004. Enhancement of carbon sequestration in US soils. BioScience, 54:895–908.

Post W.M., Kwon K.C., 2000. Soil carbon sequestration and land-use change: processes and potential. Global Change Biol., 6:317–327.

Powlson D.S., Bhogal A., Chambers B.J., Coleman K., Macdonald A.J., Goulding K.W.T., Whitmore A.P., 2012. The potential to increase soil carbon stocks through reduced tillage or

organic material additions in England and Wales: a case study. Agric. Ecosyst. Environ., 146:23–33.

Pribyl D.W., 2010. A critical review of the conventional SOC to SOM conversion factor. Geoderma, 156:75–83.

Rabenarivo M., Chapuis-Lardy L., Brunet D., Chotte J., Rabeharisoa L., Barthès B.G., 2013. Comparing near and mid-infrared reflectance spectroscopy for determining properties of Malagasy soils, using global or LOCAL calibration. J. Near Infrared Spectrosc., 21:495–509.

Riefolo C., Castrignanò A., Colombo C., Conforti M., Ruggieri S., Vitti C., Buttafuoco G., 2019. Investigation of soil surface organic and inorganic carbon contents in a low-intensity farming system using laboratory visible and near-infrared spectroscopy. Archives of Agronomy and Soil Science, 66(10), pp. 1436–1448.

Rienzi E.A., Mijatovic B., Mueller T.G., Matocha C.J., Sikora F.J., Castrignanò A., 2014. Prediction of soil organic carbon under varying moisture levels using reflectance spectroscopy. Soil Science Society of America Journal, 78:958–967.

Rinnan Å., van den Berg F., Engelsen S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry, 28(10):1201-1222.

Robinson D.A., Campbell C.S., Hopmanns J.W., Hornbuckle B.K., Jones S.B., Knight R., Wendroth O., 2008. Soil moisture measurement for ecological and hydrological watre-shade scale observatories: A review. Vadose Zone Journal, 7(1):358-389.

Rouse J.W., Haas R.H., Scheel J.A., Deering D.W., 1973. Monitoring Vegetation Systems in the Great Plains with ERTS. In Proceedings of the 3rd Earth Resource Technology Satellite (ERTS) Symposium,Washington, DC, USA, pp. 48–62.

Ryabchykov O., Guo S., Bocklitz T., 2018. Analyzing Raman spectroscopic data. Physical Sciences Reviews. 2018; 20170043.

Sahoo U.K., Singh S.L., Gogoi A., Kenye A., Sahoo S.S., 2019. Active and passive soil organic carbon pools as affected by different land use types in Mizoram, Northeast India. PLoS ONE, 14(7):e0219969.

Santi C., Certini G., D'Acqui L.P., 2006. Direct determination of organic carbon by dry combustion in soils with carbonates. Commun. Soil Sci. Plant Anal., 37:155–162.

Sarkhot D.V., Grunwald S., Ge Y., Morgan C.L.S., 2011. Comparison and detection of total and available soil carbon fractions using visible/near infrared diffuse reflectance spectroscopy. Geoderma, 164:22-32.

Savitzky A., Golay M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least-Squares Procedures. Analytical Chemistry, 36:1627-1639.

Schnitzer M., Schulten H. R., 1992. The analysis of soil organic matter by pyrolysis- field ionization mass spectrometry. Soil Science Society of America Journal, 56:1811-1817.

Schmidt M.W.I., Torn M.S., Abiven S., Dittmar T., Guggenberger G., Janssen I.A., Kleber M., Kögel-Knaber I., Lehmann J., Manning D.A.C., Nannipieri P., Rasse D.P., Weiner S., Trumbore S.E., 2011. Persistence of soil organic matter as an ecosystem property. Nature, 478:49-56.

Schillaci C., Acutis M., Lombardo L., Lipani A., Fantappiè M., Märker M., Saia S., 2017. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. Science of The Total Environment, 601-602:821-832.

Sequi P., 2010. Fondamenti di chimica del suolo. Pàtron editore, Bologna.

Shivangi S.S., Kumari M., 2020. Comparative analysis of different vegetation indices with respect to atmospheric particulate pollution using sentinel data. Applied Computing and Geosciences, 7:100032.

Silvero N.E.Q., Demattê J.A.M., Amorim M.T.A., dos Santos N.V., Safanelli J.L., Poppiel R.R., Mendes W.d.S., Bonfatti B.R., 2021. Soil variability and quantification based on Sentinel-2 and Landsat-8 bare soil images: A comparison. Remote Sensing of Environment, 252:112117.

Six J., Bossuyt H., Degryze S., Denef K., 2004. A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. Soil Tillage Res., 79:7–31.

Smith P., 2008. Land use change and soil organic carbon dynamics. Nutr. Cycl. Agroecosyst., 81:169–178.

Soil Science Division Staff, 2017. Soil survey manual. C. Ditzler, K. Scheffe, and H.C. Monger (eds.). USDA Handbook 18. Government Printing Office, Washington, D.C.

Springer U., Klee J., 1954. Pru¨fung der leistungsfa¨higkeit von einigen wichtigen verfahren zur bestimmung des kohlenstoffs mittels chromschwefelsaure sowie vorschlag einer neuen schnellmethode. J. Plant Nutr. Soil Sci. 64:1–26.

Stenberg B., 2010. Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. Geoderma, 158:15-22.

Stenberg B., Viscarra Rossel R.A., Mouazen A.M., Wetterlind J., 2010. Visible and Near Infrared Spectroscopy in Soil Science. In Donald L. Sparks, editor: Advances in Agronomy, Vol. 107, Burlington: Academic Press, pp. 163-215.

Stevenson FJ. 1994. Humus chemistry: genesis, composition, reactions. 2nd ed. New York: Wiley.

Stockmann U., Adams M.A., Crawford J.W., Field D.J., Henakaarchchi N., Jenkins M., Minasny B., McBratney A.B., de Remy de Courcelles V., Singh K., Wheeler I., Abbott L., Angers D.A., Baldock J., Bird M., Brookes P.C., Chenu C., Jastrowh J.D., Lal R., Lehmann J., O'Donnell A.G., Parton W.J., Whitehead D., Zimmermann M., 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agriculture, Ecosystems and Environment, 164:80-99.

Swift R.S., 1996. Organic matter characterization. In: Sparks DL, editor. Methods of soil analysis. Part 3. Chemical methods. SSSA Book Series no. 5. Madison, WI: Soil science Society of America and American Society Agronomy. p 1011-69.

Tate R.L., 1987. Soil organic matter: Biological and ecological effects. New York: John Wiley. 291 p.

Thaler E.A., Larsen I.J., Yu Q., 2019. A New Index for Remote Sensing of Soil Organic Carbon Based Solely on Visible Wavelengths. Soil Sci. Soc. Am. J., 83:1443–1450.

Tokarski D., Wiesmeier M., Weissmannovà H.D., Kalbitz K., Demyan M.S., Kučerìk J., Siewert C., 2020. Linking thermogravimetric data with soil organic carbon fractions. Geoderma, 362:114124.

Torrent J., Barrón V., 2008. Diffuse Reflectance Spectroscopy. Soil Science Society of America, 667: 367-383.

Trumbore S., 2009. Radiocarbon and soil carbon dynamics. Ann. Rev. Earth Planet. Sci., 37:47–66.

Vasquez G.M., Grunwald S., Sickman J.O., 2009. Modeling of Soil Organic Carbon Fractions Using Visible–Near-Infrared Spectroscopy. Soil Sci. Soc. Am. J., 73:176-184.

Vapnik V., 2000. The Nature of Statistical Learning Theory, Second. Springer, New York, USA.

Vaudour E., Gomez C., Loiseau T., Baghdadi N., Loubet B., Arrouays D., Ali L., Lagacherie P., 2019. The Impact of Acquisition Date on the Prediction Performance of Topsoil Organic Carbon from Sentinel-2 for Croplands. Remote Sensing, 11:2143.

Vezzani L., Ghisetti F., Festa A., 2004. Geological map of Molise (scale 1:100000). S.E.L.CA.

Viscarra Rossel R.A., 2008. ParLeS: software for chemometric analysis of spectroscopic data. Chemom Intell Lab Syst., 90:72–83.

Viscarra Rossel R.A., Behrens T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma, 158:46–54.

Viscarra Rossel R.A., Hicks W.S., 2015. Soil organic carbon and its fractions estimated by visible–near infrared transfer functions. European Journal of Soil Science, 66:438–450.

Viscarra Rossel R.A., Walvoort T.D., McBratney A.B., Janik L.J., Skjemstad J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma, 131:59-75.

von Lützow M., Kögel-Knabner I., Ekschmitt K., Flessa H., Guggenberger G., Matzner E., Marschner B., 2007. SOM fractionation methods: Relevance to functional pools and to stabilization mechanisms. Soil Biol. Biochem., 39:2183–2207.

von Lützow M., Kögel-Knabner I., Ekschmitt K., Matzner E., Guggenberger G., Marschner B., Flessa H., 2006. Stabilization of organic matter in temperate soils: Mechanisms and their relevance under different soil conditions – a review. European Journal of Soil Science, 57:426-445.

von Lützow M., Kogel-Knabner, I. 2009. Temperature sensitivity of soil organic matter decomposition-what do we know? Biology and Fertility of Soils, 46(1):1-15.

VV.AA., 2006. Metodi di valutazione dei suoli e delle terre. E. Costantini Coordinator, Cantagalli edition.

Walkley A., Black I.A., 1934. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. Soil Science, 37:29-38.

Webster R., Oliver M.A., 2007. Geostatistics for Environmental Scientists, Statistics in Practice. John Wiley & Sons, Ltd, Chichester, UK.

Wen D., He N., Zhang J., 2016. Dynamics of Soil Organic Carbon and Aggregate Stability with Grazing Exclusion in the Inner Mongolian Grasslands. PLoS ONE, 11(1):e0146757.

Wiesenberg G.L.B., Dorodnikov M., Kuzyakov Y., 2010. Source determination of lipids in bulk soil and soil density fractions after four years of wheat cropping. Geodarma, 156:267–277.

Wold S., Martens H., Wold H., 1983. The multivariate calibration method in chemistry solved by the PLS method. In: Ruhe A., Kagstrom B. (eds.), Lecture Notes in Mathematics. Proceedings of the Conference on Matrix Pencils. Springer-Verlag, Heidelberg. pp 286-293.

Wold S., Sjöström M., Eriksson L., 2001. PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst., 58:109–130.

Workman J.J., Weyer L., 2008. Practical Guide to Interpretive Near-Infrared Spectroscopy. 2nd ed. Boca Raton (FL): CRC Press.

Wösten J.H.M., Lilly A., Nemes A., Bas C.L., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma, 90:169–185.

Yang F., Tian J., Fang H., Gao Y., Xu M., Lou Y., Zhou B., Kuzyakov Y., 2019. Functional soil organic matter fractions, microbial community, and enzyme activities in a mollisol under 35 years manure and mineral fertilization. J Soil Sci Plant Nutr, 19:430–439.

Zhang L., Chen X., Xu Y., Jin M., Ye X., Gao H., Chu W., Mai J., Thompson M.L., 2020. Soil labile organic carbon fractions and soil enzyme activities after 10 years of continuous fertilization and wheat residue incorporation. Scientific Reports, 10:11318.

Zhou T., Geng Y., Ji C., Xu X., Wang H., Pan J., Bumberger J., Haase D., Laush A., 2021. Prediction of soil organic carbon and the C:N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images. Science of the Total Environment, 755:142661.

Zimmermann M., Leifeld J., Schmidt M.W.I., Smith P., Fuhrer J., 2007. Measured soil organic matter fractions can be related to pools in the RothC model. Eur. J. Soil Sci., 58:658–667.

Zornoza R., Guerrero C., Mataix-Solera C., Scow K.M., Arcenegui V., Mataix-Beneyto J., 2008. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. Soil Biology & Biochemistry, 40:1923–1930.

Zovko M., Romić D., Colombo C., Di Iorio E., Romić M., Buttafuoco G., Castrignanò A., 2018. A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia. Geoderma, 332:60-72.