UNIVERSITÀ 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 XXXI)

Related disciplinary scientific section: AGR/14 (Pedologia)

PhD thesis

NEW APPLICATIONS OF VISNIR SPECTROSCOPY FOR THE PREDICTION OF SOIL PROPERTIES

Coordinator of curriculum: Prof. Giuseppe Maiorano

Supervisor: Prof. Claudio Colombo

Co-Supervisor: Prof. Edoardo A. C. Costantini

PhD Student: Romina Lorenzetti

155901

ACADEMIC YEAR 2017/2018

Summary

Soil is known to be an unrenewable limited natural resource. All the ecosystemic functions that it absolves are required for the survival of humanity, such as agricultural and forestal products, regulation of water movement in the landscape, sink for Carbon in the frame of climate change, environmental filter for pollutants, space for infrastructure and recreation.

Preserving these functionalities requires to preserve needs an appropriated knowledge and interpretation of the information about soils. Since the biggest limits to data availability are by the spendable amounts of time and money, researches are currently focused on new faster and cheaper instruments of investigation. By the time that multivariate statistical methodologies and computer-based analysis took widely place, visible (Vis) (400–780 nm) and near-infrared reflectance (NIR) (780–2500 nm) (VisNIR) spectroscopy seemed one of the fair solutions for this needing. VisNIR spectroscopy is recognized to be a non-destructive, fast, inexpensive and precise alternative to wet chemistry, enable to obtain a large amount of data partly losing the reliability of traditional analyzes, due to the inevitable error that occur in predicted data.

Several studies were carried out primarily to test the predictive power of VisNIR spectroscopy, and several data was been produced by the last three decades. Nevertheless, some potentialities and limits of VisNIR soil spectroscopy are still a little unexplored and new questions may occur as a result of the huge amount of available data.

This thesis embodies a collection of novel studies related to the use of multivariate information provided by VisNIR spectroscopy. As a whole they experienced questions: at different reference scales (national and local); with different investigation method (multivariate calibration and multivariate classification); in different field of application (taxonomy, monitoring, scientific purpose). The idea was to focus on some of the main issues that are emerging at the state of the art of knowledge about VisNIR spectroscopy in soil science.

Briefly, Chapter 1 contains the object of study, focusing on what is already known and what may need to be further discussed. Thus, the first part is about the soil that is the object of investigation and the second in about the VisNIR spectroscopy.

Chapter 2 deals with the possibility to explore VisNIR potentiality over some still poorly explored properties. A case study for taxonomic purpose was carried out over volcanic soils. The results of this chapter show the close connection between soil spectroscopy attribute variation and soil taxonomic units linked to volcanic properties.

Chapter 3 deals with a national library used at local scale for organic carbon (OC) monitoring. It highlights the increasing economic advantage of NIR with the decrease of OC change.

Chapter 4 focuses on the possibility of using the raw information contained in the VisNIR region of the spectra to characterize the soils rather than using individual soil attributes. In this chapter, VisNIR showed a higher power in detaching slight differences due to effect of a very recent changing in the forest management respect to wet biochemical analyses.

Chapter 5 focuses on the status of the large spectral library in Italy. To my knowledge, this PhD represents the first study that included the building of an Italian National Spectral Library.

Finally, Chapter 6 discusses the main findings of this thesis.

Riassunto

Il suolo è una risorsa naturale limitata e non rinnovabile e le funzioni ecosistemiche da esso assolte sono indispensabili alla sopravvivenza dell'uomo. Tra queste: le produzioni agricole e forestali, la regolazione del movimento dei corpi idrici, lo stoccaggio del carbonio nella cornice del cambiamento climatico, l'azione filtrante per le sostanze inquinanti, e infine superficie per le infrastrutture e le attività ricreative.

Preservate tali funzionalità richiede una profonda e dettagliata conoscenza della risorsa suolo e un adeguata interpretazione delle informazioni. Dal momento che diponibilità di tempo e disponibilità economiche rappresentano i maggiori limiti alla raccolta di dati relativi alla risorsa suolo, le ricerche sono attualmente incentrate su nuovi strumenti di indagine più veloci e meno costosi. Con la diffusione di metodologie statistiche multivariate e l'analisi computerizzata, la spettroscopia visibile (Vis) (400-780 nm) e nel vicino infrarosso (NIR) (780-2500 nm) (VisNIR) ha iniziato a prendere piede come valida alternativa alla chimica tradizionale. La spettroscopia VisNIR è riconosciuta come uno strumento non distruttivo, veloce, economico e preciso, consente di ottenere una grande quantità di dati, perdendo in parte l'affidabilità delle analisi tradizionali, a causa dell'inevitabile errore che si verifica nei modelli di stima.

Negli ultimi tre decenni sono stati condotti diversi studi principalmente per testare la potenza predittiva della spettroscopia VisNIR e sono stati raggiunti numerosi risultati. Tuttavia, alcune potenzialità e limiti della spettroscopia del suolo VisNIR rimangono ancora poco esplorate e in virtù dell'ingente quantità di informazione che si sta producendo, nuovi interrogativi potrebbero insorgere nella comunità scientifica.

Questa tesi comprende una raccolta di studi su tematiche attuali e innovative relative all'uso della spettroscopia VisNIR in abito pedologico. Nel complesso le attività di ricerca sono state svolte: a diverse scale di riferimento (nazionali e locali); con diversi metodi di indagine (calibrazione multivariata e classificazione multivariata); in diversi campi di applicazione (tassonomia, monitoraggio, finalità scientifica). L'idea era di affrontare alcuni dei principali problemi che stanno emergendo allo stato dell'arte dell'uso della spettroscopia VisNIR nella scienza del suolo.

In breve, il Capitolo 1 contiene l'oggetto di studio, riassumendo ciò che è già noto e ciò che potrebbe essere necessario discutere ulteriormente. Quindi, la prima parte riguarda il suolo che è l'oggetto di indagine e il secondo la spettroscopia VisNIR.

Il capitolo 2 riguarda le potenzialità del VisNIR nella predizione di alcune proprietà ancora poco esplorate, nello specifico: le proprietà vulcaniche dei suoli. Rappresenta un caso di studio di uso del VIsNIR per scopi tassonomici. I risultati di questo capitolo mostrano la stretta

connessione tra la variazione dell'attributo della spettroscopia del suolo e le unità tassonomiche del suolo legate alle proprietà vulcaniche.

Il capitolo 3 viene affrontato il tema dell'uso di una libreria spettrale nazionale a livello locale per il monitoraggio del carbonio organico (OC), con l'obiettivo di evidenziare per quali range di variabilità e differenze attese da monitorare, l'uso del Vis NIR è da preferisti alla chimica tradizionale.

Il capitolo 4 si concentra sulla possibilità di tracciare la qualità relative di un suolo, utilizzando le informazioni grezze contenute nella regione VisNIR degli spettri piuttosto che i singoli attributi biochimici del suolo. Sono stati indagati campioni di suolo relativi ad aree boschive avviate a differenti tipi di gestioni forestali Il VisNIR ha mostrato una maggiore potenza nel l'evidenziare le lievi differenze dovute all'effetto di un cambiamento molto recente dei trattamenti, rispetto alle analisi biochimiche tradizionali.

Il capitolo 5 riguarda la costruzione di un'ampia libreria spettrale dei suoli in Italia. Per quanto è a mia conoscenza, questo dottorato di ricerca rappresenta il primo tentativo di costruzione di una biblioteca nazionale spettrale dei suoli italiani.

Acknowledgements

This period as a PhD student represented the possibility to introduce a high formative parenthesis in my working career, since it was arrived in my life after more than 10 years from my master thesis and after about 6 years of research activity at the research Center CREA-AA. For this reason, I firstly wish to thank my colleagues which I had by my side before and during my PhD, collaborating in the most suitable way: Simone Priori, Maria Fantappiè, Roberto Barbetti, Giovanni L'Abate, and Lorenzo D'Avino, thanks for sharing with me achievements, expectations and defeat. I' m proud to belong to this working stuff up to now. I want especially to thank Edoardo Costantini for believing in me, and for embodying a great example of love for scientific working.

A big thank you is also for Cesar Guerrero, my tutor in the foreign university, for his generosity in teaching me whatever he could.

At the end, I need to say many thanks to the people closest to me. They have been my indispensable energy to carry out this project in the frame of a so insecure job, and my courage to overcome the misgivings, looking to the future.

Publications

Colombo C., Lorenzetti R., Di Iorio E., Costantini E. A. C., 2017. Stima delle proprietà andiche tramite spettroscopia diffusa nel dominio visnir su suoli vulcanici europei. 42° Congresso Nazionale della Società Italiana della Scienza del Suolo. FIRENZE 5-7 dicembre 2017

Lorenzetti R, Guerrero C., Di Iorio E., Agnelli E.A., Colombo C., Lagomarsino A., 2018. Discriminating the effects of forest management on litter and soil in a medioterranean pine forest by VisNIR. In: ATTI DELLA ESSC Conference "Soil and Water Security: challenges for the next 30 years!". Imola, 6-8/06/2018. (BEST POSTER AWARD session 2).

Submitted pubblications

Lorenzetti R, Guerrero C., Minimum detectable difference: using composites or using NIR spectroscopy? Congress Pedometrics, 2-6 June 2019 (poster)

Erika Di Iorio , , Luana Circelli, , Romina Lorenzetti, Edoardo A.C. Costantini, Sara Perl Egendorf, Claudio Colombo. Estimation of andic properties from Vis-NIR diffuse reflectance spectroscopy for volcanic soil classification. Catena

Alessandra Lagomarsino; Alessandro Elio Agnelli; Gianluigi Mazza; Romina Lorenzetti; Claudio Colombo; Caterina Bartoli; Carlo Viti; Roberta Pastorelli. "Litter fractions and dynamics in a degraded pine forest after thinning treatments". European Journal of Forest Research (EFOR).

TABLE OF CONTENTS

List of figures

List of tables

INTRODUCTION

Scientists agree that land – and particularly arable land – is a limited resource. Land is required for very different activities and purposes: it is needed for agricultural and forestall production of food, fiber, and energy, it is consumed to build villages, cities, industrial production plants, and all the infrastructures required for our style of living, it is demanded as space for leisure and recreation, and even used to get rid of any type of waste. The land has an imporant role in the regulation of water movement in the landscape, as environmental filter for metals, nutrients, and other contaminants, as biological habitat and gene reserve (Stenberg et al., 2010). The ability of a soil to support any of these functions depends on its structure, composition, and chemical, biological, and physical properties; and all of them are all both spatially and temporally variable (Blum, 1993; Bouma, 1997; Harris et al., 1996; Jenny, 1980; Karlen et al., 1997). The necessity to maintain these soil functionalities for this and the future generations, necessarily implies a sustainable use of the resource.

Sustainable use inevitably depends on a land-use planning and soil monitoring, based on an appropriated knowledge and interpretation of the information about soils, their characteristics, their dynamics and of course their geographic spatializations. The biggest limits to these data availability deal with the spendable amounts of time and money. Thus, in addition to traditional soil mapping and classical chemical analyses, supplementary approaches are growing their importance in order to use the available resources more efficiently. They are based on the possibility to estimate the target parameter by the use of predictions based on input data which can be investigated with a lower effort. They include geostatistical methods, remote sensing, and novel, fast and inexpensive analysis techniques. Among the authors, it is opened up this idea that Visible and Near Infrared spectroscopy can be considered as promising technique to generate a big amount of soil data due to the efficiency in reducing time and cost efforts. Some recent studies deal with quantifying time and cost savings by means of VisNIR (McBratney et al., 2006; Nocita et al., 2015). However further investigations should be carried out in order to define clearly the limits within VisNIR spectroscopy shows to outperform traditional wet chemistry approach, taking into account not only the costs of the analysis but also the effect of the error in estimation induces on the usefulness of the data.

Spectral soil properties have a cumulative nature that depends on the combination of soil components: organic matter and water content, but also from physical aspect as the soil particle size and their organization. For this reason, the spectral properties have the potentiality for being use as predictive information for estimation of both qualitative and quantitative soil parameters, and the field of its applicability in the soil science was being growing along the decades.

Numerous soil parameters were investigated on the base of VisNIR spectroscopy during the last years: it was often used to estimate SOC and its fractions as it usually achieved very good results. Many other properties were investigated, like clay content, iron oxides, cation exchange capacities, moisture content, several minerals (Hunt, 1977), as well as soil biological

properties like biomass and respiration rate (Viscarra Rossel et al., 2006; Stenberg et al., 2010), achieving different degrees of success.

Nevertless, the relations between spectral properties and some diagnostic soil properties are still poorly investigated, as for instance, the properties linked to volcanic soils. This kind of information could get stronger the success of VisNIR spectroscopy exploitation for taxonomic issue.

In the last years, a new goal of soil spectroscopy regards the possibility to assess soil variability and changes in soil conditions based directly on spectral data. This approach is based on the fact that spectral soil properties allow a holistic approach to soil, as many physical, chemical, and biological parameters contribute to soil reflectance. Therefore, estimating and comparing soil quality could be raise skipping the intermediary step of estimating soil properties. (Odlare et al., 2005; Islam et al., 2005; Dematte et al., 2004). Indeed, McBratney et al. (2006) suggested to estimate pedotransfer functions directly by spectral data.

Due to the fast and cheap acquisition of spectral data, both local and national and global databases have been growing, even available for free (e.g. LUCAS project (Toth et al., 2013)). The construction of large libraries has been promoted from the researchers to increase the predictability of VisNIR spectroscopy (Brown et al., 2005; Shepherd 2002).

Nocita et al. (2015) recently presented an extensive review of the state of the art of the general library. Among the Global libraries, the ICRAF-ISRIC world soil spectral library collected 4438 samples from Africa, Asia, Europe, North America, and South America; Viscarra Rossel and Webster (2012) described a large library of 21,500 Vis-NIR spectra from Australian continent; a spectral library covering the United States counts 144,833 Vis-NIR spectral signatures, (USDA, 2013); the European spectral library LUCAS consists of about 20,000 topsoil samples, collected from all over Europe (Stevens et al., 2013). Moreover, several national and regional soil spectral libraries have also been created.

As, to my knowledge, Italy doesn't own a national spectral library up to now, one of the objectives of this PhD was to collect spectra on the national territory and create a national spectral library. In line with recommendations in Nocita et al. (2015), the purpose was to start the building of the large library from scanning existing soil archives, reducing the need for costly sampling campaigns and assuring the reproducibility of analytical and spectral information.

Unfortunately, the lack of a standard for the collection of laboratory soil spectra (Nocita et al., 2015), represents at the moment an obstacle for sharing easily spectral libraries. Currently, the great increasing of spectral data availability does not run parallel to the possibility of their integrated use. Thus, a current question may be approached: what a national spectral library may add respect to a library available for free at global scale, in the frame of pedological investigations.

Given that, in this study it was proposed to address novel issues of soil spectroscopy in VisNIR domain, in agreement with what was stated till now. It aims to produce novel knowledge that could improve the use of VisNIR spectroscopy in soil science in order to increase the soil basic knowledge and the monitoring activities, which are both indispensable for a sustainable management of soil resource.

Thus, the specific focuses of the present work dealt with the following proposals:

- an application of soil spectroscopy for taxonomic purposed, focused on the classification of soil with volcanic features: the classification of soils represents a very expensive and time consuming practice, but it is a very powerful knowledge, since it summarize a wide range of information and can be a valid support for a right management of the soil resource.
- an estimation of the condition for a profitable use of soil spectroscopy with respect of wet chemistry, for monitoring change in soil organic Carbon, one of the most important and well predicted soil properties: the change in soil organic carbon content is currently one of the most important topic in the frame both of agricultural issue for farm productions, than of climatic change for the estimation of soil carbon stock capability.
- a comparison between biochemical properties and NIR spectra in checking the effects of little land use change: increasing the potential of monitoring the effect of a land use change by VisNIR spectroscopy may represent a support in the decision making process for the best management choices in the silvicoltural and agro systems.
- the creation of a Italian National Spectral Library since the lack of a large spectral library describing the Italian soil variability and a comparison with free available dataset: a new national library may represent the possibility to introduce big improvement in the knowledge of the soil properties of the soil both at national and local scale (by appropriate calibrations), according to the ability of the library to describe all the soil variability in the Italian territory.

CAP.1

SUBJECT OF STUDY

1. 1 Components and characteristics of soil

The spectroscopic investigations in this thesis had as objective the characterizations of the soil. The soil is one of the three materials covering the surface of the Earth, together with rocks and water. It results from the weathering of geological materials and death biomass (Blum et al., 2017).

According to Soil Taxonomy (1999) the soil is "a natural body comprised of solids (minerals and organic matter), liquid, and gases that occurs on the land surface and is characterized by one or both of the following: horizons, or layers, that are distinguishable from the initial material as a result of additions, losses, transfers, and transformations of energy and matter or the ability to support rooted plants in a natural environment….. Soil consists of horizons near the Earth's surface that, in contrast to the underlying parent material, have been altered by the interactions of climate, relief, and living organisms over time". Soil represents a crucial compartment as it is the basis of food production, a key player in climate and water regulation, a pool of biodiversity, a reactor to degrade contaminants. For this reason the concept of soil security is recently growing its importance (Koch et al. 2013), and it is clear the needing of more information and data to identify and describe soils and its functioning. This is why new ways of acquiring data are so important.

The soil is always an important component in the system comprising the lithosphere, the atmosphere and the biosphere. Soil properties reflect the varying nature of the interactions within this system (Rowell, 2014). Four basic components make up the soils, in a proportion depending on the environmental context. The mineral component is the dominant component in most of soils (with exception for organic soils) and it includes primary and secondary minerals, amorphous components, and water-soluble salts; the organic component is made up of dead biomass, the product of its decomposition and little living animals of the pedofauna; water and air are the other components. Soil constituents are not just mixed together, but form an organized soil body (pedon) of definite structure and distinctive chemical and physical properties. These properties determine the soil responses to the environment and to soil use (Blum et al., 2017). In this paragraph, a brief description of the soil constituents and properties was reported.

1.1.1 Mineral component of the soil

The minerals of the soil derive from the rocks which constitute the parent material. The primary minerals are incorporated into the soil when the rocks disintegrate by weathering. The primary minerals are decomposed into secondary minerals both through the chemical alteration, both through recrystallization of products. The most abundant primary mineral in the soil is Quartz, due to its high resistance to chemical alteration.

The most abundant secondary minerals are the clay minerals (phyllosilicates) and the oxy-hydroxides of Iron, Aluminum, Silicon and Titanium. The processes of alteration also produce carbonates, sulphates and phosphates. These secondary minerals are more stable in the soil the primary ones as they originate at the same temperatures and pressures conditions of the pedosphere. The secondary minerals represent the most abundant component of clay soil fraction. They are crystalline phyllosilicate (Jackson et al., 1981). Kaolinites, smectites and illite are the most common. Other less abundant clay minerals are Chlorite and Vermiculite.

1.1.2 The organic component of the soil

The soil organic matter (SOM) is a complex and heterogeneous set of organic, living and non-living components (excluding–macrofauna and mesofauna) present in the soil. The SOM includes compounds which differ for chemical and physical composition, functions and dynamics They are the result of accumulation, degradation, decomposition and resynthesise of residues, carried out by microbial organisms, animals and plants living in the soil. SOM represent the source of plant nutrients, in particular N and also S and P. SOM has also an effect on water and air content, the temperature in the soil, and soil properties as cation exchange, fertility and structure.

SOM includes both humic and non-humic components. Humus is defined as a brown to black complex variable of carbon containing compounds not recognized under a light microscope as possessing cellular organization in the form of plant and animal bodies. Humus is separated from the non-humic substances such as carbohydrates (a major fraction of soil carbon), fats, waxes, alkanes, peptides, amino acids, proteins, lipids and organic acids by the fact that distinct chemical formulae can be written for these non humic substances. Most small molecules of non humic substances are rapidly degraded by microorganisms within the soil. Conversely, soil humus is slow to decompose (degrade) under natural soil conditions.

The components of humus are the Humic substances. They are Humic Acids, Fulvic Acids, and Humin. (Zanella et al., 2001). The Humic Acids are the major humic component of soil and although water insoluble under very acidic conditions (pH<2), are readily soluble at higher pH's. They are dark brown to black in colour. The Fulvic Acids are water soluble in all pH conditions, and they are yellow to yellow-brown in colour. The Humin is the fraction of humic substances that is not soluble in water at any pH value and in alkali. Humins are black in colour.

Humus substances can be bound by various clay minerals through some linkages known as clay-humus complex. The complexes between humus substances and clays are mainly formed by bridging through the exchangeable cations like Ca, Mg and Al, so that humus substances penetrate into the inter layer space of the crystalline clay mineral lattice. The responsible for this process are mainly the earthworms by digestion, but also other animals like diplopods, chilopods, isopods. The clay-humus complex has both a direct and indirect effect over some important soil properties, including: the increase of soil ability to retain chemical elements and the improvement of the formation of stable aggregates.

The scientific literature of the 15 last years show that humic substances contribute little to total SOC, and are mostly an operational SOC fraction.

A very important fraction of soil organic matter is represented by the soil enzymes, even if quantitatively they are a minority, since all biochemical actions are dependent upon, or related to them.

Enzymes are proteins specialized in the catalysis of biological reactions. Although microorganisms are the primary source of soil enzymes, plants and animals also contribute to the soil enzyme pool. Soil enzymes are important for catalysing innumerable reaction necessary for life processes of micro-organisms in soils, decomposition of organic residues, cycling of nutrients and formation of organic matter and soil structure.

The extracellular portion of soil enzymes cam be stabilized and inactivated by humic substances by the formation of covalent bonding (Stevenson & Cole, 1999). The stabilization makes these enzymes less subject to microbial degradation and less active up to stop their function. However, many of these bonds are relatively weak during periods of pH change within the soil, these enzymes can be released.

Soil enzymes respond rapidly to any changes in soil management practices and environmental conditions. Their activities are closely related to physio-chemical and bio-logical properties of the soil. Hence, soil enzymes are used as sensors for soil microbial status, for soil physio-chemical conditions, and for the influence of soil treatments or climatic factors on soil fertility (Rao et al., 2014).

1.1.3 Water and air in the soil

The water and the air contained into the soil fill in a vicarious way the space between the particles and the aggregates of the solid materials (Blum et al., 2017). The air composition differs from the atmospheric one for a higher concentration of CO_2, because of the effect of the respiration of roots and pedofauna. The air in the soil in very important for allowing respiration under the surface and for the controlling role in the oxidation and reduction processes of the pedogenesis.

The water in the soil is a solution containing dissolved salts. A constant exchange of ions occurs between the solid phase of the soil and the liquid phase and between the latter and the roots of the plant. The content in water allows to make water and nutrient in the solution available for the plant and it is a driving force in the soil forming processes (Blum et al., 2017).

The water content is the balance between the input by precipitation and from groundwater, and the water lost for the canopy interception, evapotranspiration, surface runoff, and percolation.

Water is retained in the pores of the soil from the forces of capillary attraction. Because of polar nature of water molecules, soil solids attract water molecules. Soil water can be divided into categories, for descriptive purposes, based on its tension. The water that it is drained quickly from the ground after saturation is called gravitational water. Generally, it occupies larger pores at a pressure of 0.1-0.2 bar. The water capillary is, instead, retained in the smaller pores from

capillary forces of attraction, at a pressure between 0.1-31 bar. It represents the main source for the growth of vegetation. Water adsorbed strongly from the solids of the soil it is at a pressure above 31 bar it represents hygroscopic water (Brady, 1989).

1.1.4 Main chemical soil properties

Soil chemistry is the interaction of various chemical constituents that takes place among soil particles and in the soil solution—the water retained by soil.

The nutrient-holding capacity is a very important chemical property for the ecological function of the soil. Soils can hold onto nutritional elements thank to the adsorption of cations on the part of the negative charges on the soil particles. The ability to hold cation nutrients is called the cation-exchange capacity (CEC) and is an important characteristic of soils in that it relates to the ability to retain nutrients and prevent nutrient leaching (Curtaz and Zanini, 2012). Clay retains more nutrients than coarser soils, because of the greater surface area. A sand may have a CEC of under 10, a very low value. Any CEC above 50 is high, and such soils should be able to hold ample nutrients. Nutrient availability varies markedly according to pH. The best pH for overall nutrient availability is around 6.5, which is one reason why this is an optimal pH for most plants. In low-pH soils many cation-exchange sites are occupied by H^+ cations and they cannot hold other cations. Thus, if cations are not held by particles, they can leach out of the soil. Therefore, low-pH soils are more likely to be deficient in nutrients such as Magnesium, Calcium or Potassium (Curtaz and Zanini, 2012).

Along with ion exchange properties, two other important indices of the soil chemical environment are the reaction and redox potential.

Soil pH (or soil reaction), is a measure of the number of hydrogen ions (H^+) present in a solution. The pH scale runs from 0 to 14 and it is a measure of alkalinity and acidity of the water solution into the soil. Soil pH ranges between 3 and 9, usually 3-7 in temperate and humid climate condition and 6-9 in warm conditions th et allow strong evapotranspiration.

The parent material of soils initially influences soil pH. For instance, granitic soils are acidic and limestone-based soils are alkaline. However, soil pH can change over time under the action of: rainfall, root growth and decay of organic matter by soil microorganisms, fertilization and irrigation. Changes in soil pH, whether natural processes or human activities cause them, occur slowly thanks to the high resistance of soil to change in pH (buffering capacity) (Curtaz and Zanini, 2012). The highest buffering capacities belongs to soils high in clay or organic matter (high CECs). Calcareous soils often have high buffering capacities as lime can neutralize acid.

The redox potential (Eh) indicates the tendency of a soil to be reduced or oxidized. Redox reactions are very important in soil genesis. Reduction, as a chemical process, occurs when an atom accepts an electron. This process increases the valence of an anion or decreases the valence

of a cation. Oxidation is the reverse process and occurs when an atom loses an electron. Oxidation–reduction (redox) reactions in soils are mainly controlled by microbial activity and the presence of a supply of carbon for the microbes. When the supply of oxygen for the respiration is terminated, as is the case under conditions of saturation, the microbial activities switch from aerobic to facultative and eventually to anaerobic respiration. Soils that remain saturated (gley) are chronically reduced (Fiedler et al., 2007).

The redox potential (Eh) is expressed as a function of the potential referred to the reaction of H^+ in standard condition of temperature and pressure that is the lowest value of a reduction occurring in a water solution. Thus, it is high if the redox system has a high oxidation power (Curtaz and Zanini, 2012). Eh in soils generally ranges between −1 and +1 V. This variation is due to the buffering effect or poise of water on redox reactions (O2/H2O; H2O/H2) (Bartlett and James, 1993, Bartlett and James, 1995). Changes of external conditions, such as precipitation and water table, temperature, and availability of organic matter, can all lead to changes in Eh values. Consequently, the redox potential can vary by several orders of magnitude both temporally and spatially (Gao et al., 2002, Vepraskas et al., 1999).

Salinity is also a feature belonging to the chemical properties. It concerns the of soil that refers to the amount of salts in the soil. It can be estimated by measuring the electrical conductivity (EC) of an extracted soil solution. Since low rainfall prevents leaching of salts, some soils, particularly in arid regions, may hold high levels of salt. Also clay soils are more prone to salt accumulation. Some fertilizers and amendments also can increase salinity. High level of salinity can affect plant growth in several ways, directly and indirectly: an high osmotic potential of the soil solution makes water absorption very difficult for the plant that has to use more energy; there may be some ion-specific toxicity; An imbalance in the salts content may result in a competition between elements leading to an interference with uptake of essential nutrients.

1.1.5. Main physical soil properties

Physical properties play an important role in determining soil suitability for agricultural, environmental and engineering uses. The retention availability, and movement of water and nutrients to plants, the easiness in penetration of roots, the air flow are directly associated with physical properties of the soil. Physical properties also influence the chemical and biological properties (Phogat et al., 2015). Here is a short discussion on the most relevant physical properties for plant growth.

Soil texture refers to the prominent size range of mineral particles<2 mm. The groups of different size range of mineral particles are known as textural fractions, namely: sand, silt and clay. Soil texture is the relative proportion of sand, silt and clay content on weight basis. There are three broad primary textural groups of soils (sandy, loamy and clayey) and twelve textural

classes (Soil Survey Staff, 2006). It is more or less a static property affecting almost all other soil properties (Phogat et al., 2015).

The soil particles in natural conditions are bonded together into larger units, called aggregates.

Soil structure is defined as the arrangement and organization of soil particles in the soil, and the tendency of individual soil particles to bind together in aggregates. The structure development is influenced by the amount and type of clay and the its binding with organic compounds; the amount and type of organic matter, since it provides food for soil fungi and bacteria and their secretion of cementing agents; the presence of iron and aluminum oxides as cementing agents; roots act as holding soil together, and protects soil surface. The aggregation creates intra-aggregates and inter-aggregate pore space, thereby changing flow paths for water, gases, solutes and pollutants. Water supply, aeration, availability of plant nutrients, heat, root penetration, microbial activity, and other soil characteristics and functions are strongly affected by the soil structure. Strong aggregation decreases detachability and transportability of soil particles by water or wind and thus, reduces runoff and soil erosion. It is affected by tillage, cultivation and application of fertilizers, manures, lime, gypsum and irrigation (Phogat et al., 2015). Four basic organizations summarize the types of soil structure: single grains; massive; granular aggregates, aggregates formed by segregation process such as drying and shrinking of clay minerals (angular and subangular blocky, play) (Blum et at., 2017)

The system of empty space kept clear by particles represents the soil porosity. The soil pore space between particles of soil in heathy soil are large and plentiful enough to retain the water, oxygen and nutrients that plants need to absorb through their roots. Soil porosity falls into one of three categories: micro-pores, macro-pores or bio-pores. These three categories perform different functions with respect to soil permeability and water holding capacity. For instance, water and nutrients in macro-pores will be lost to gravity more quickly, while the very small spaces of micro-pores are not as affected by gravity and retain water and nutrients longer. Soil porosity is affected by soil particle texture, soil structure, soil compaction and quantity of organic material. Silt and clay soils have a finer texture and sub-micro porosity th et allow to retain more water than coarse, sandy soils, which have larger macro-pores. Both finely textured soils with micro-pores and coarse soil with macro-pores may also contain large voids known as bio-pores. Bio-pores are the spaces between soil particles created by earthworms, other insects or decaying plant roots.

Bulk density (BD) is the density (mass per unit volume) of a dry soil including the pore space (Blum et al. 2017). The BD is influenced by texture, structure, moisture content, organic matter and management practices of soil. In coarse textured soils, it varies from 1.40 to 1.75 Mg m-3 while in fine textured soils, it normally ranges from 1.10 to 1.40 Mg m-3. The BD decreases

with increase in organic matter content and fineness of soil texture. Higher values of BD indicate more compactness of the soil (Phogat et al., 2015).

Another physical aspect of the soil regards its dynamic properties: when certain forces act on a body, the forces do not produce any bodily motion, but produce a relative displacement of particles of the body resulting in a change of shape or size or both of the body (Phogat et al., 2015). The most important dynamic properties for agriculture and ecosistemic soil function include soil consistency, crusting, compaction, and permeability. Consistency includes the strength of a soil to withstand an applied stress (rupture resistance); the rate of change and the physical condition soil attains when subjected to compression (manner of failure); the capacity of soil to adhere to other objects (stickiness), the degree to which reworked soil can be permanently deformed without rupturing (plasticity); the ability of soil in a confined state to resist penetration by a rigid object (penetration resistance). All of them depend from the state of cohesion and adhesion of the soil mass (Blum, 2017).

Soil crust is a thin compacted surface layer of higher bulk density than the soil immediately beneath which is formed due to dispersion of soil aggregates as a result of wetting and impact of rain drops, and its subsequent rapid drying. The thickness of crust may vary from mm to few cm depending upon the amount and type of clay, and silt content of the soil. Soils having organic matter less than 1% are more prone to crusting (Phogat et al., 2015).

Soil compaction is the process of increasing bulk density and reducing pore volume as a result of the applied pressure. It leads to destruction of larger pores, re-arrangement of solid particles and compression of air within the pore spaces in the soil. The degree of compaction depends upon the nature of clay minerals, type of exchangeable cations, water content and extent of manipulation of the soil. A compacted layer is commonly found just below the usually tilled layer of soil. The compacted layer often restricts root penetration and reduces water and nutrient uptake by crops (Phogat et al., 2015).

Permeability expresses the easiness with which soil allows fluid to pass through.

Soil colour is a physical property that gives an indication about the state of the pedological processes and the type of minerals in the soil. Under oxidized conditions (well-drainage) the red colour occurs, due to the abundance of iron oxide in the soil; the accumulation of highly decayed organic matter lead to due to dark colour; the presence of hydrated iron oxides and hydroxide give yellow colour; colours of soil matrix and mottles are indicative of the water and drainage conditions in the soil and hence suitability of the soil for aquaculture (Blum et al., 2017).

1.1.6 Soil quality

Safety of environmental condition as well all the soil functions are to be preserved for future generations (Reeves, 1997), and maintaining or improving soil quality is crucial.

Together with water and air quality, soil quality forms the environmental quality (Andrews et al., 2002). Differently from water and air quality that are limited to the degree of soil pollution (Carter et al., 1997; Davidson, 2000), soil quality has a broader definition. According to Doran & Parkin (1994, 1996), it is "The capacity of a soil to function within ecosystem and land-use boundaries to sustain biological productivity, maintain environmental quality, and promote plant and animal health".

The great soil complexity leads to the need of defining soil quality with respect to the desired function (Giacometti et al., 2013). Thus, several soil physical, chemical and biological properties have been proposed for useful indicators of soil quality (Reeves, 1997; Arshad and Martin, 2002; Anderson, 2003; Schloter et al., 2003; Winding et al., 2005). Due to its influence over the other soil properties, SOM has been suggested as the most important single soil quality indicator (Giacometti et al., 2013). However, when soil quality has to be estimated in the frame of monitoring activity, it has to be taken into account the slow process of SOM changes induced by land use management, so to require a long time before being experimentally detectable (Rasmussen and Collins, 1991; Bending et al., 2004; Körschens, 2006). Moreover, biogeochemical mechanisms behind the observed modifications in SOM are not explained only by the SOM.

Since the soil microbiota mediates most of the soil processes (Dick, 1992) and it quickly adapts to environmental constrains by adjusting its biomass, activity rates and community composition (Schloter et al., 2003), microbial biomass and soil enzyme activities had been successfully applied as Soil quality indicators to measure the impact of land management on soil (e.g. Pajares, et al., 2009).

For a very exhaustive review over the soil quality indicators it is acltually possible refer to the recent publication by Bünemannet al. (2018).

1. 2 The Visible and Near Infrared Spectroscopy

1.2.1.1 Principles of Visible and Near Infrared Spectroscopy

One of the most important results of the quantum theory is th et all atoms and molecules can only be found in states with specific and characteristic values of energy. The central theme of spectroscopy is the study of the transitions between two different energy states of a system of atoms or molecules. The transitions take place with absorption or emission of energy in the form of electromagnetic radiation having a frequency n (or wavelength λ) given by Bohr relation:

$$\Delta E = hn = hc/\lambda \qquad [Eq\ 1.1]$$

where ΔE is the variation of energy of a system, h the Planck constant, and c the light speed.
Thus, the smaller is the variation of energy among the levels, the bigger the wavelength. The wavelength range of VisNIR is reported in the figure 1.1.

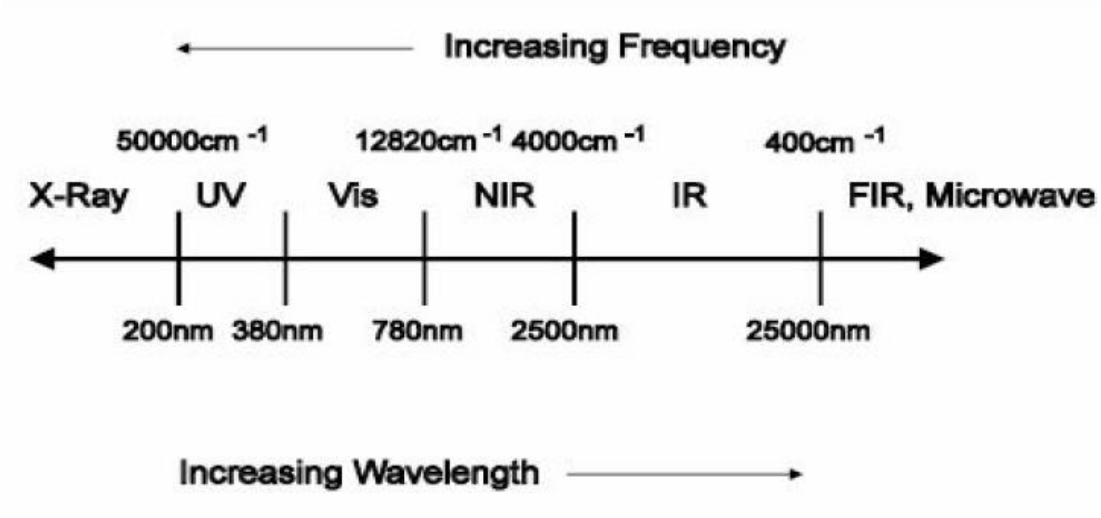


Figure 1.1: Range of Electromagnetic radiation (Aenugu et al., 2011).

The transitions in the energy levels lead to variations in the energy state of an electron (electronic absorption) or in the movements of atomic nuclei (transition). These last movements can be transitional, rotational or vibrational. When an electromagnetic wave hits the soil it mainly generates vibrational movements and change in the energy state of an electron, while in many soil materials the molecular and rotational transitions are limited (Irons et al., 1989; Drury, 1993).

In the VisNIR spectroscopy the relevant energetic transitions are the electronic adsorption and the vibrational transition (Hunt,1977).

The Visible region is characterized by the electronic absorptions caused by the movement of electrons from one orbit to a higher-energy orbit (Næs et al., 2002). The spectral response in the Vis region is not very strong (Nocita et al., 2015). Neverthless, some authors had derived quantitative information from the spectra (Viscarra Rossel et al., 2006, Owen, 2000).

In the NIR the chemical adsorption mechanisms mostly consist in vibrational energy transitions into molecules. Indeed, these last typically require energy of a frequency that corresponds to the IR region of the electromagnetic spectrum. The vibrational movements consist in oscillations of an atom with respect to another, within the molecule. There are two types of vibrational movements: the molecular bond stretching and the inter-bond angles bending. The vibrational energy has quantized levels. A molecule can move from a vibrational energy level to a higher one by the absorption of photons of infrared radiation. The energetic transition between the fundamental level of energy and the first excited level is called fundamental vibration. Transitions may also occur between the fundamental level and second excited level or the following ones (overtones). The intensity of these successive abortions will decrease by a factor between of 10 and 100 (Næs et al., 2002). In particular, the NIR region shows overtone and combination modes of the fundamental atoms vibrations in molecules that

are active in mid- and far-infrared (Nocita et al., 2015). They may be liked to a specific bound or to the entire molecular structure.

Combination bands that appear between 1900 nm and 2500 nm are the result of vibrational interactions. For instance, their frequencies are the sums of multiples of each interacting frequency (Dyrby et al., 2002). The two or more fundamental vibrations sharing their absorption in a combination, can be observed separately in the MIR (middle infrared) range (tab.1.1). As a consequence of the combinations, in the NIR spectrum the absorptions occur at unexpected positions, and the regions of adsorption occur as broad peaks caused by the overlapping of multitude of different absorptions (Næs et al., 2002) (fig. 1.2).

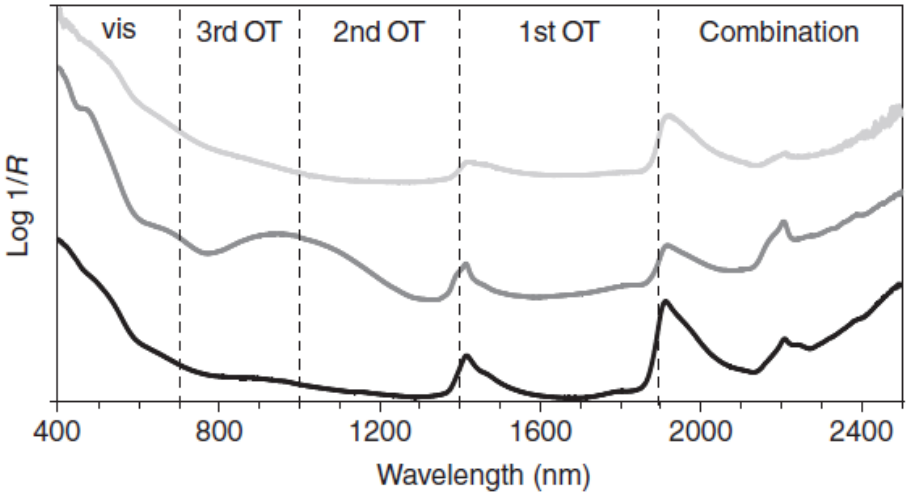


Figure 1.2: Soil spectra in VisNIR domain (400-2500 nm), showing approximately where the combination, first, second, and third overtone vibrations occur (Stenberg et al., 2010).

Table 1.1: Characteristics of Visible and NIR domains

	Visible	NIR
Wavelength (nm)	400-760	760-2500
Wavenumber (cm^{-1})	25000-13200	13200-4000
Overtones and combination		X
Electronic process	X	X
Absorption s related to	electrons	CH/OC/NH functionalities

1.2.1.2 The reflectance.

When an electromagnetic wave hits a material, it can be transmitted, reflected and absorbed. If the material is opaque, the transmittance cannot occur. This is the case of soil, where a part of the electromagnetic energy is reflected, and the other part is absorbed. The proportions

of absorption and reflection are wavelength dependent and provide information about the physical and chemical properties of the material (Clark, 1999).

The reflectance has two different forms: the specular reflectance and diffuse reflectance. The first one refers to the reflection occurring with the same angle of incidence. The diffuse reflectance refers to the reflectance that occurs at all other angles. It is expressed by the ratio of the intensity of the light reflected in all the directions and the intensity of the light of incidence perpendicularly on the reflecting object (Bédidi and Cervelle, 1993).

Diffuse reflectance is dominant when the particle size (or the roughness) is bigger than the ratio between wavelength and angle of incidence (Cervelle and Flay, 1995). If the material is opaque, like the soil, specular reflectance is very low (Siesler, 2008).

The total absorption of a specific material comprises a physical and a chemical component.

The physical component is mainly caused by the geometry of the measured surface (particle size, shape of particles,….), whereas the chemical component is related to the composition of the material.

For the measure of diffuse reflectance, a spectrophotometer is used. It is composed by a sphere with a white cover insight; a font of radiations and a detector for recording the reflected radiation. The intensity of the reflected radiation is collected in correspondence of a little pore where the sensor detector is located (figure 1.3).

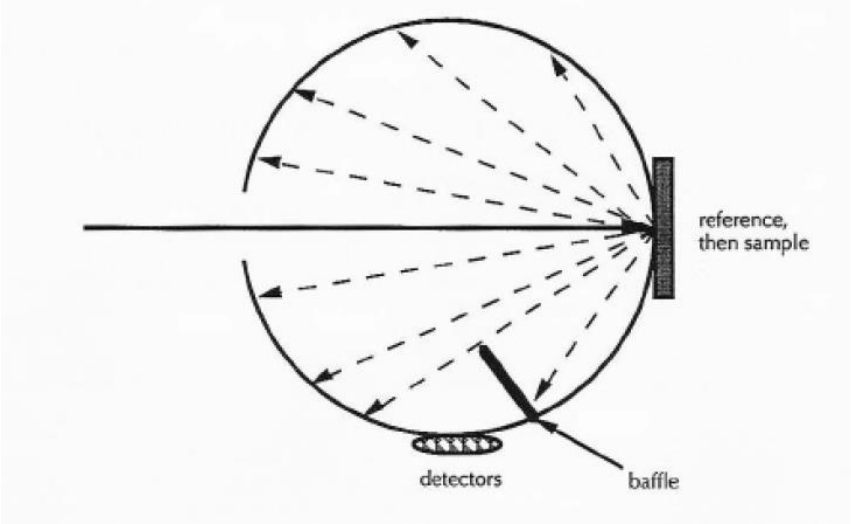


Figure 1.3: Schematic representation of an integrating sphere used to obtain a diffuse reflectance spectrum (modified by Springsteen, A. (1994)).

The frequencies at which light is absorbed appear as a reduced signal of reflected radiation and are displayed in % reflectance (R), which can then be transformed to apparent absorbance (A):

$$A=\log(1/R) \qquad \text{[Eq.1.2]}$$

The reflectance is usually referred to a white standard material with known reflectance. Spectralon (politetrafluoretilene) is one of the most adopted with reference surface, since it is very stable along the time, but also barrio sulfate or magnesium monoxide.

1.2.2 Soil Spectral signature signatures in the VisNIR domain.

A spectral signature of soil in the VisNIR domain is characterized by its general form, the intensity of the signal, and the specific adoption bands. These last depend on the electronic and vibrational transition that we have already mentioned. Both the mineral, organic and water components affect the soil spectrum. Several authors summarized the spectral properties of soil components.

1.2.2.1 The mineral component.

A huge work was published by Hunt, Salisbury et al., already in 1970-1976, about spectral properties of mineral and rocks. More recent reviews on soil spectral properties were carried out by Madeira Netto (1995) and Leone (2000).

The purity, the elemental composition and the crystalline structure of the primary minerals are all factors that determine the reflectance of the soil. The use of spectroscopy for the study of Quartz in the soil is rather complicated. Framework silicates such as quartz do not have prominent absorption features in the UV–vis–NIR region. Their intense fundamental vibrations occur in the mid infrared around 10,000 nm (Nguyen et al., 1991). Hunt (1977) found that small absorption bands occur near 850 nm, 1200 nm, 1400 nm, 1900 nm due to the vibrational combinations and overtones of molecular water contained in various locations in the mineral. Moreover, the isomorphic substitutions that may occur, lead to different reflective properties (Hunt and Salisbury, 1970).

The presence of clay minerals such as kaolinite [$Al_2Si_2O_5(OH)_4$], montmorillonite [$Na0.33(Al_{1..67}Mg_{0.33})$ $Si_4O_{10}(OH)_2$]. and illite [$(Al_2(Si_{3.85}Al_{0.15})O_{10}(OH)_2)$], induces an absorption band near 2200 and 2300 nm because of the combination of vibrations associated with the OH bond and the OH-Al-OH bonds (Hunt et al.,1971 Chabrillat et al., 2002) The metal-OH bend plus O–H stretch combination near 2200 nm and 2300 nm are recognized to be diagnostic absorption features in clay mineral identification (Clark et al., 1990). The absorption features of kaolin, smectite and illite are similar. Specifically, Kaolinites and smectites differ for the presence of a band at 1900 nm. Viscarra et al., (2006) found that the spectrum of montmorillonite (the high-alumina member of the group of Smectites) is that of typical of a water-rich mineral with an intense absorption peaks at 1400 nm band which may be attributed to the first overtone of the O–H stretch, and also at 1900 nm band which is due to the combinations of the H–O–H bend with the O–H stretches. On the opposite site the spectrum of kaolinite only has a peak at 1400 nm band, indicating that only O–H is present, and only a small amount of water is included in the structure (Clark et al., 1990, Hunt and Salisbury, 1970, Goetz et al.,

2001). The absorption features of illite are the same but less well defined, and it has hydroxyl bands at 1400 nm and between 2200 and 2600 nm (Hunt and Salisbury, 1970).

Other clay minerals have specific spectral responses in NIR spectral domain, relative to vibrations associated with Mg and Fe-OH bonds.

Goethite (FeOOH) and ematite (Fe2O3) are the most frequent ferric (Fe^{3+}) minerals in the soils. They result from Fe contained into the primary minerals as Fe^{2+}, by the oxidation due to the alteration processes. Because of the electronic process, Iron in Fe^{3+} status produces three absorbing bands between 400 and 1000 nm. Hematite has its typical absorption bands at 550 ,630 and 920 nm, while goethite at 480, 650 and 920 nm (Hunt et al.,1971, Morris et al., 1985).

Inorganic carbon consists primarily of the two minerals calcite and dolomite. These minerals have distinct absorption features at 2500 to 2550 nm, 2300 to 2350 nm, 2120 to 2160 nm, 1970 to 2000 nm, and 1850 to 1870 nm (Clark et al., 1990; Hunt and Salisbury 1971; Gaffey 1986). The spectral responses of these minerals are related to the combinations of the four possible vibrational transition movements of the planar ions CO_3^2 (Hunt and Salisbury, 1971; Clark et al., 1990): symmetrical stretch of C-O bond at 9407 nm, out-of-plane bend at 11400 nm, symmetrical stretch at 14150 nm and in-plane bend at a 14700 nm. The position of the bands moves according to the carbonate composition.

The soil spectra also contain information about soluble salts which can affected salinity and alkalinity of the soils. Their spectral signatures were summarized by several authors (Hunt and Salisbury, 1971; Hunt *et al.,* 1972; Mulders, 1987; Mougenot *et al.,* 1993; van der Meer, 1995, Mougenot *et al.*, 1993). Salts have specific spectral responses especially in NIR spectral domain, except for halite (Mougenot *et al.,* 1993). This last (NaCl) is completely transparent like the quartz and its formula and cubic structure give any response nor NIR neither in Vis domains (Hunt *et al.,* 1972: Eastes, 1989). Only water bands can be highlighted at 1400 and 1900 nm, due to the eventual humidity of the samples (Hunt et al., 1972; Mulders, 1987; Mougenot *et al.,* 1993). The gypsum ($CaSO_4 \cdot 2H_2O$) absorption bands are between 1000 and 2500 nm (the biggest ones at 1450, 1750, 1900 and 2200 nm) (Hunt and Salisbury, 1971; Mulders, 1987) due to the combination of fundamental mode of the water molecules in the crystalline lattice. Some spectral signatures of the main mineral components of the soil are reported in the following figure (fig. 1.4)

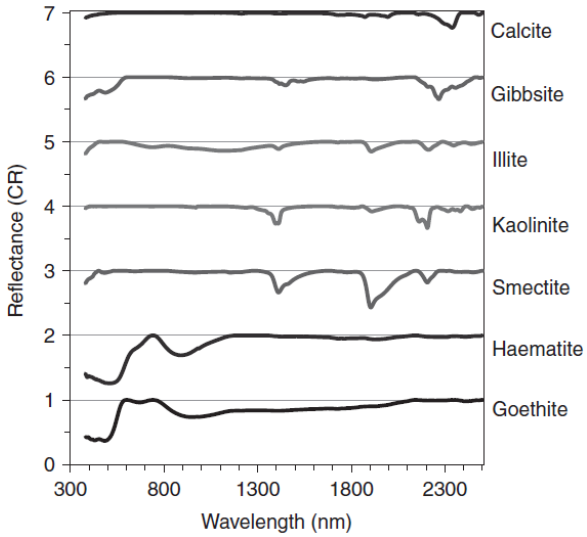


Figure 1.4: Spectral signatures of the most common soil minerals. Reflectance was augmented 1 unit for each signature. (Stenberg et al., 2010).

1.2.2.2 The organic component.

The active bonds in organic carbon are O-H, C-N, N-H, and C=O groups, which are primarily in the mid infrared region (Malley et al., 2002). The overtones and combinations of the hydroxyl bonds are located in the NIR region.

Soil organic matter content and the composition of organic constituents are known to have a strong influence on soil reflectance long ago (e.g. Baumgardner et al., 1985). The clearest effect of soil organic carbon on the soil spectra signature are that the soil organic matter content generates a soil reflectance decrease (Baumgardner et al., 1985), and that most of the solar energy absorbed by humic materials is between 300 and 500 nm (Gaffney et al., 1996).

Very different spectral properties belong to the humic and fulvic acids (Obukhov and Orlov, 1964; Henderson et al., 1992). The first ones have a very low reflectance (<2%), due to the presence of several components with high absorbance in the Vis domain (e.g. phenolic constituents and their oxidation product, amino acids and their condensation products (Flaig et al., 1975). On the other hands, the fulvic acid can reach a reflectance of 20% at 750 nm (Obukhov and Orlov, 1964).

Despite of the several studies about the soil organic matter (Shields et al., 1968; Karmonov and Rozhkov, 1972; Vinogradov, 1981, Henderson et al., 1992) the effects of the different components of the organic matter are still not completely unveiled. Recent studies found an acceptable predictability of fulvic and humic components by VisNIR spectroscopy (e.g. Vergnoux et al., 2009).

Due to the wide importance given to the enzymatic activity in motoring soil quality, a certain amount of studies were carried out in the last years for its estimation by spectral soil properties (Zonroza et al., 2009, Dick et al., 2013) The enzymatic activity belongs to that group of soil properties that are often shown to be accurately predicted by some studies, but not by

others (as for pH, CEC, nutrients, etc). The explanation for this is suggested to the lack of direct relationships between the spectra and soil properties (Stenberg et al., 2010).

1.2.2.3 The water content of soil

The polar water molecules of water are characterized by three vibrational transitions: O-H symmetrical stretch; O–H–O symmetrical bend, O–H asymmetrical stretch. They produce fundamental bands 3106, 6080 and 2903 nm, when water is at the liquid status (Hunt and Salisbury, 1970; Clark *et al*., 1990).

The water content effect on soil spectral properties has been studied since long time (Bowers and Hanks, 1965; Shields *et al.*, 1968; Cierniewski, 1985; Celis-Custer, 1980; Bedidì *et al.*, 1992). It is known a general reflectance decrease with the increase of water content in the soil (Bowers and Hanks, 1965). When water occurs in the mineral, these bands have a combined effect in the soil spectra: two bands appear at 1400 nm due to the O-H stretching and at 1900 nm due to bending plus stretching).

The presence of both the bends highlights that molecules of water are in the minerals of soil. Differently, when only 1400 peak appears it refers to the –OH group of some mineral structure (Hunt and Salisbury, 1970; Clark *et al.*, 1990). A parameter capturing the relative extension of the water absorption band near 1940 nm was found most useful to estimate soil water content (Bowers and Smit, 1965).

When water is present in the soil, it may obscure spectral information because of the O-H bonds at 1400 and 1900 nm which are important spectral signatures of clay minerals (Bricklemeyer et al., 2010). Anyway, recent study demonstrated that heterogeneous water content did not affect the clay prediction accuracy, but it did affect IC and OC prediction accuracy (Waiser et al., 2007). An important step forward has been the recent investigations on the possibility to take into account the effect of a heterogeneous moisture conditions into predictive models for soil properties. (e.g. Nocita et al. 2009, Minasny et al., 2011). These studies allow a qualitative use of VisNIR spectroscopy directly on filed conditions.

The following table (table 1.2) summarizes the most important VisNIR adsorption bands observed for soil samples.

Table 1.2: Important VisNIR absorption bands in soil constituents (Hunt, 1977).

Soil constituents	Wavelength (nm)	Explanation
Water (H_2O)	1400-1500	Combination of symmetric and asymmetric OH-stretch
	1900-2000	Combination of H-O-H bend with asymmetric OH-stretch
	2200-2800	The absorption maxima of fundamentals symmetric and asymmetric OH-stretch fall into MIR range (2600-2800 nm), only a decline in reflectance is

		visible in NIR range beyond 2200 nm
Hydroxyl group (OH)	1400-1500	Combination of symmetric and asymmetric OH-stretch
	2200-2400	OH attached to metal, paired absorption peak, exact position depends on metal and structure of mineral.
Carbonates	1900, 2000, 2160, 2350, 2550	Combination and overtones of C-O bonds.
Quartz (SiO_2)	None	Si-O bonds possess strong absorptions in the MIR but not in VisNIR domain

1.2.2.4 Soil physical properties.

The soil signature in VisNIR domain is strongly influenced by the particle size. It was observed an inverse correlation between reflectance and particle size due to the fact that smaller particles are more densely packed, leading to higher concentration of the measured material and lower porosity (Bowers and Hanks, 1965). A finer particle size also generates smoother surfaces reducing the possibility to lose part of the irradiation in the surface roughness (Baumgardner et al.,1985; Bowers and Hanks, 1965). However, they observed that the particle size effect is not the same. Most of the reflectance increase was observed for particles smaller 0.4mm, while only slight differences were observed between fractions of bigger particles. Stenberg et al. (2002) and successively Stenberg (2010) studied the integration of sandy fraction over the predictability of SOC content in the soil. The authors observed that predictions of SOC had larger errors when the soils contained larger amounts of sand and that the errors in the sandiest soils were clearly dominated by over estimations.

Likewise the particle size, the soil aggregation of undisturbed samples influence the spectral response: clayey soils have a lower reflectance, as they have a higher tendency to create bigger aggregates than those of sandy soils (Myers and Allen, 1968).

1.2.2.5 Soil quality

It is known that in the last decades degradation processes are affecting the soils, progressively actions have been taken in order to evaluate and reduce the loss of soil quality.

Assessing soil quality required an integrated consideration of key soil properties Anyway the selection of monitoring variables is considered difficult and their interpretation affected by subjective evaluations (Cècillon et al., 2009a). Several systems for the estimation of soil quality involve many soil analyses, so that monitoring remains very expensive especially at regional and national scale. By contrast, VisNIR spectroscopy is a fast and cost efficient technique, and, most of all, spectral information depend on several soil properties, which we may considered summarized in the spectral signature of a soil for that reason some authors have recently suggested the use of VisNIR as integrated measure of soil quality. Some studies focused on the

prediction by spectral information of soil parameters chosen as soil quality indicator. Palmborg and Nordgren (1996) have shown that measurements of organic matter quality as seen by NIR can be used in evaluations of effects of moderate heavy metal contamination on microbial processes in forested sites of northern Sweden. At field scale, Sudduth et al. (2009) evaluated the ability of visible and near infrared (VNIR) spectroscopy to estimate soil quality indicators, such as carbon content of particulate organic matter, nitrogen content of particulate organic matter, permanganate oxidizable C, and other properties of soil fertility (P, K, pH). soil organic carbon, Nitrogen. But they generally obtained poor predictions, since small range of variability occurred both in spectral properties and in soil quality across the test area.

Other studies suggested the use of principal components of PCA for discriminating clusters of soil samples differing in their quality, since they synthesize information on soil condition. Velasquez et al. (2005) and Zonroza et al. (2009) verified the possibility to use the principal component of PCA and discriminant analysis to separate soils from different land uses. Cécillon et al. (2009b) employed a NIR spectral data set of Mediterranean topsoils and earthworm casts collected in areas affected by wildfire. They successfully separated Soil samples and biogenic structures by PCA on NIR spectra, depicting the influence of earthworms on soil quality, as previously demonstrated by Hedde et al. (2005).

Cohen et al. (2006) evaluated indicators of minimally/moderately/severely degraded ecological conditions based on biogeochemical soil properties and VisNIR spectroscopy, by combining ordinal logistic regression and classification trees of soil NIR spectra. While spectra were comparably effective at discriminating minimally degraded sites, they were significantly more effective at discriminating severely degraded sites.

As suggested in the paper review by Cècillon et al. (2009a), soil spectroscopy is a promising tool for soil quality assessment, “as reliable quantification of particular soil functions, ecosystem services or threats can be evaluated from a flight campaign or a simple NIR scanning of a soil sample”.

1.3 Extracting information from soil spectra in VisNIR domain.

As explained in the previous sections, NIR spectra are very information-rich due to the number of overlapping absorption bands (Blanco and Villarroya, 2002). Differently from other region of electromagnetic spectrum where each peach is directly connected to the presence of specific molecular bond, in the interpretation of VisNIR spectra it is often difficult in practice to find selective wavelengths for the chemical constituents in the samples. The absorbances et all wavelengths are affected by a number of the chemical and physical properties of the sample. The question has been solved in spectroscopy by the use of a multivariate statistics approach.

Thus, the adopted chemometric methods for extracting the necessary information from VisNIR spectra are multivariate calibrations and multivariate classifications.

1.3.1 Multivariate calibration

The multivariate calibration is probably the chemometric methodology which has attracted the most interest so far (Naes et al. ,2002). It is used for quantitative spectral analysis by means

of sophisticated statistical techniques to discern the response of soil attributes from spectral characteristics. Various methods have been used for quantitative spectral analysis. Some example is reported by Viscarra et al., (2006):

- multiple regression analysis (MRA) to relate specific bands in the NIR to a number of soil properties (Ben-Dor and Banin, 1995).
- stepwise multiple linear regression (SMLR) for the estimation of various soil properties from the NIR spectra of soil acquired by a field-deployed on-the-go soil sensing system (Shibusawa et al.., 2001).
- multivariate adaptive regression splines (MARS) for the estimation of soil properties from soil spectral libraries. Fideñcio et al. (2002) employed radial basis function networks (RBFN) to relate soil organic matter to soil spectra in the NIR region (Shepherd and Walsh, 2002).
- artificial neural networks to estimate soil organic matter, phosphorus and potassium from the VIS–NIR spectrum. (Daniel et al., 2003)
- principal components regression (PCR) (e.g. Chang et al., 2001) and partial least squares regression (PLSR) (e.g. McCarty et al., 2002)

When extracting information from VisNIR, it has to be taken into account the problem of the collinearity: the high correlation and linear (or near) relations that occur among the spectral variables. Standard regression, for instance cannot properly work when problems of collinearity appear among the predictive variables. Thus, successful techniques have to be able to manage collinearity. In chemometric literature two approach are most popular: removing co-linear predictive variables from the models and transforming the variables into their linear combinations to be used in the models (Naes et al., 2002).

Sometimes, also the presence of non-linearity between predictors and predicted parameters may represent an obstacle. The non-linearity may exist between the target variable on one side and each of the spectral variables on the other (univariate non-linearity) but can disappear in the multivariate relation, often, without any consequences in a multivariate linear approach. When non-linear relation exists between the target variable and all the spectral variables simultaneously (multivariate non-linearity) it may be a limit for the multivariate linear approach. In some cases, no-linear methods can give substantial improvements. Nevertless, linear prediction techniques are usually able to work well (Naes et al., 2002, Naes et al., 1990), as different methods can be adopted to solve non-linearity. An overview of them is summarized in the following scheme (fig. 1.5).

In the soil science literature PLSR regression is one of the most common analysis techniques (Goetz et al., 2001; Dunn et al., 2002; Waiser et al., 2007; Morgan et al., 2009; Viscarra-Rossel et al., 2009). It performs similar to PCR, but in a slightly different manner. PCR decomposes the spectra into a set of eigenvectors and scores and in a second moment the regression with soil attributes is performed. PLSR actually uses the soil information during the decomposition process. PLSR takes advantage of the correlation that exists between the spectra and the soil, thus the resulting spectral vectors are directly related to the soil attribute (Geladi and

Kowalski, 1986). The advantages of PLSR are that it handles collinearity, it is robust in terms of data noise and missing values, and unlike PCR it balances the two objectives of explaining response and predictor variation (thus calibrations and predictions are more robust) and it performs the decomposition and regression in a single step (Viscarra et al., 2006). The majority of work comparing PLSR regression to BRT and MLR show that PLSR regression performs just as well or better than the other two methods (Brown et al., 2005).

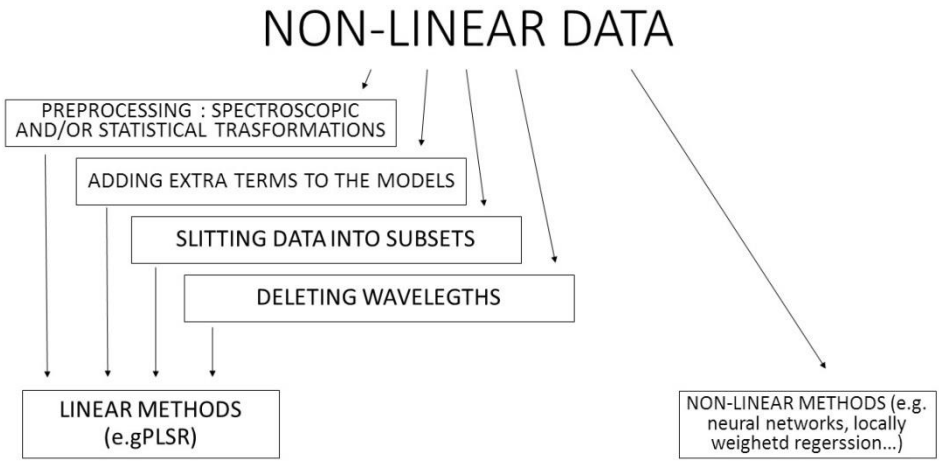


Figure 1.5: Overview of solutions for non-linearity (modified by Naes et al., 2002).

The first formal work using spectroscopy techniques as we see today was most probably the one of Bowers and Hanks (1965), where they observed the influence of moisture, organic matter and particle size on the absorbance of the wavelengths ranging from 500 to 2500 nm.

The first studies, in the 70–80's, focused on the interpretation and classification of soil spectra (e.g., Stoner & Baumgardner in 1981 collected a spectral library containing representative soil samples of the United States, and identified five typical spectral curves corresponding to five soil classes). Later, soil spectroscopy started to adopt a quantitative approach to predict many soil properties. Thus, we found studies about organic Carbon (Ben-Dor & Banin, 1995, Krishnan et al., 1980), moisture (Dalal and Henry, 1986), and so on. The use of multivariate statistical approach allowed to test the use of soil spectroscopy an alternative to traditional methods of wet chemistry to measure common soil properties. Many studies followed on same quantitative approach to predict several soil properties such as organic carbon (Gobrecht et al., 2014), texture (Sorensen and Dalsgaard, 2005), cationic exchange capacity (CEC) (Cañasveras Sànchez et al., 2012), total phosphorus (P) (Abdi et al., 2012), exchangeable potassium (K) (He et al., 2005), and electrical conductivity (Ben Dor et al., 2002; Viscarra Rossel et al., 2006; Todorova et al., 2011). Currently, some studies have presented extensively reviews of soil spectroscopic models published in the literature. The work of Soriano-Disla et al. (2014) was one of the last one and it contains the list of soil properties that could be determined

by means of diffuse reflectance (Nocita et al 2015). According to Viscarra Rossel et al., (2006) calibrations for total and organic Carbon are probably the most frequent, followed by clay content. These two, together with total N, are also those with the best chance of success since they have well-recognized absorption features in the vis–NIR region. Some other frequently reported properties include pH, extractable P, K, Fe, Ca, Na, Mg, and CEC, as well as properties that are dependent on combinations of other soil properties, such as lime requirement and mineralisable N. Results for these are typically moderate and often highly variable.

The promising results combined with the easy and fast operations for collecting VisNIR spectra data, are leading the current spectroscopy to rapidly move toward on-the-field studies (Hedley et al., 2014; Ackerson et al., 2015). VisNIR spectroscopy on-the-field is one of the proximal soil sensing that mostly paved the way for the development of the "digital soil morphometrics", a new soil science sub-discipline, the objective of which is the application of tools and techniques for measuring and quantifying soil profile attributes and deriving continuous depth functions (Hartemink and Minasny, 2014).

1.3.2 Multivariate classification

Beside the multivariate calibrations, the multivariate classifications complete the most important chemometric methodologies for extracting information by VisNIR spectra. Two types of analyses can be performed.: the cluster analysis and the discriminant analysis (Naes et al., 2002, Blanco and Villarroya, 2002) The first one has usually an explorative function and it is performed to check the presence of groups (unsupervised classification); discriminant analysis is used to create classification rules which are able to allocate a sample into one of the already know groups (supervised classification). Differently from multivariate calibration, the discriminant analysis can be used to predict categorical variable or/and which class a sample belong. It is possible to find examples of these applications in the recent soil spectroscopy literature. Many of these studies focused on taxonomic purposes. Soil taxonomic systems are built considering soil properties and their organization in the soil profile. Therefore, it seems logical that taxonomic group could be correlated to the soil spectral properties as they result by the combinations of several soil chemical and physical properties. Mouazen et al. (2005) found the possibility to discriminate four texture groups by on the field soil spectra. Mouazen et al., (2006) suggested that a factorial discriminant analysis on VisNIR data collected in the field can be used to correctly classify soils into groups of different soil water content level, particularly when soil variability is very low. Ben-Dor et al. (2008) and Demattê et al. (2012), worked on the possibility to create spectrally derived diagnostic horizons. They observed that commonly described soil morphological horizons had a distinctive "spectral signature" in the Vis-NIR domain, highlighting its possible use in discrimination and further classification. Vasques et al. (2014) observed the strong relation between depth and spectral behavior and investigated on the potentiality of VisNIR spectra collected from multiples depths, to the soil classification

Some studies tested the effect of creating subsets based on VisNIR properties, on the predictability of multivariate calibration models. For instance, McDowellet al. 2012) used VisNIR spectra to create subsets which differ for Carbon content level, soil order, and spectral

classification, looking for the possibility to outperform prediction of Soil Total Carbon with diffuse reflectance spectroscopy.

A lower number of study investigated the potential of VisNIR to directly discriminate soil variation at landscape level. Actually, this VisNIR application may be very useful for detect the occurrence of some effect in the soil induced by environment change (e.g. landuse destination, agricultural practices) and in general to indicate a general soil quality level or pedodiversity status. This potentiality of VisNIR is still very poor explored. Some suggestion over, this VisNIR application is contained in Zornoza et al. (2009). Here a Separation of land uses (agricultural, abandoned agricultural, and forests) by means of discriminant analysis was carried on sets of physical, chemical biochemical soil variables and NIR soil dataset separately. The information contained in NIR spectra outperformed the results obtained by the other type of information, suggesting the ability of NIR spectroscopy to keep up the soil variability in the landscape.

1.4 Large library

Large spectral libraries are needed to provide general and robust models over large areas that are characterized by a large soil diversity (Nocita 2015a) They are also very useful to prediction a local scale as some authors have already demonstrated (Guerrero et al. 2010) The importance of large library stays in the possibility to include into a large variability also samples similar to those that have to be predicted. Indeed, predictions need spectroscopic empirical calibrations and accurate predictions cannot be produced for samples not represented in the spectral libraries. Nocita et al. (2015a) provided a wide and exhaustive review of the state of the art.

Global Soil Spectral Library where 23,631 soil spectra have been currently collected from 92 countries in seven continents (Africa, Antarctica, Asia, Europe with 3518 of which 209 from Italy, North and Central America, Oceania, South America) (Viscarra Rossel et al., 2016).

LUCAS (Land Use/Cover Area frame Statistical Survey), is a European spectral library collecting about 20,000 topsoil (0 -20 cm) with 13 analyzed soil properties (Stevens et al., 2013).

Large libray are stated to be available also for some countries, such as s the ones for France (Gogé et al., 2012), Czech Republic (Brodsky et al., 2011), Denmark (Knadel et al., 2012), Florida (Vasques et al., 2010), and Brazil (Bellinaso et al., 2010).

1.5 Economic advantage by VisNIR Soil Spectroscopy

Actually, only few studies deal with the quantitative discussion on the Cost/Benefit Analysis of Soil Spectroscopy. For instance, O'Rourke and Holden (2011) compared Walkley-Black, total organic C analyzer, VisNIR diffuse reflectance spectroscopy, and laboratory hyperspectral imaging, in order to find the best method in terms of: costs per sample, analytical accuracy, and time for SOC analysis. They found that Vis-NIR spectroscopy and laboratory hyperspectral imaging (800 and 1720 nm) were 10 times cheaper than total organic C analyzer.

As summarize by Nocita et al (2015), for the African soil information system (AfSIS) project, the World Agroforestry Centre (ICRAF) found that the price of SOC analysis by

spectroscopy was one-third of that by a C/N analyzer. Similar results were reported by the Joint Research Centre of European Commission, for the SOC analysis of about 20,000 samples in the frame of LUCAS soil monitoring network (Toth et al., 2013).

Good results were obtained also when the objective of prediction was monitoring. At local scale, Schwartz et al. (2012) showed that the use of spectral analysis to monitor petroleum contamination in soils provided much better results at lower cost compared to certified laboratories using traditional analytical chemistry methods. Since the higher is the number of samples to be compared, the higher are the cost efforts, the economic advantage is even more important for large-scale soil assessment, where the number of measurements is very high.

Thus, for estimating the real saving, it is of great importance the evaluation of:

- what about the minimum amount of data that have to be compared for evaluating a change. This dimension depends on how much the observed data populations are different, namely the dimension of the expected change and the data variability (see MDD in cap.3);
- the costs of such amount of data by means of traditional analysis;
- the costs of such amount of data by means of predictive models. They include the costs of traditional analysis employed for the calibration of the predictive models, and the costs for the acquisition of the spectral signatures.

Indeed, if a model needs a very big calibration dataset for a high accuracy (low error), but the number of samples needed for monitoring is not very big, the economic advantage of VisNIR spectroscopy, may be low or even absent. Currently the literature is poor in study the boundaries of the conditions when VisNIR-monitoring approach can lead to a real advantage. This kind of knowledge may be helpful to decide whether or not monitoring by VisNIR spectroscopy techniques have to be preferred.

1.6 Thesis objectives

1) the possibility to explore VisNIR potential over other poorly explored diagnostic properties. This case study deals with volcanic soil classification. In order to overcome the expense of traditional soil laboratory analysis and the limitations of pedotransfer functions, this research aims to test: i) VisNIR spectroscopy in supervised chemometric classification of soils with different degrees of andicity ii) the use of VisNIR spectroscopy to estimate andic properties required for soil classification (iron and aluminum forms, allophane, phosphate retention, vitric content etc.) by different multivariate statistical approaches.

2) a contribution to the Cost/Benefit Analysis of soil organic carbon monitoring under agricultural practice changes and at local scale, by means of VisNIR spectroscopy. The main issue was to clarify the boundaries of the conditions when VisNIR-monitoring approach can lead to a real saving, taking into account the efforts necessary to balance any residual error in prediction. In particular, the study aimed i) to find what type of NIR calibration was the best one for monitoring OC in the soil; ii) to propose a quantification of the effort in terms of number size for balancing the error in checking the change; iii) to border the conditions (the range of OC

increase and variance we expected to be monitored) where NIR-based approach can assure a cheaper monitoring, compared with traditional wet chemistry analysis.

3) the possibility to use VisNIR data directly as an indicator of changing condition, bypassing the creation calibration models for the estimations of some soil parameters. The study was carried out under forest conditions, trying to check the effect of a very recent change in management practices. The spectral property ability was compared to those of biochemical properties, and enzyme activity in particular, since its very short reaction time to changed conditions.

4) to provide a National Soil Spectral library for the Italian country. The aim was to build a National Soil Spectral library following the lines suggested by recent reviews on the usefulness of large library. Secondly, a rough comparison was carried on in order to evaluate the utility of an Italian National library in the frame of the already available data for the Eropean territory.

References

Abdi, D., Tremblay, G.F., Ziadi, N., Bélanger, G., Parent, L.-E., 2012. Predicting soil phosphorus-related properties using near-infrared reflectance spectroscopy. Soil Sci. Soc. Am. J. 76, 2318e2326.

Ackerson, J.P., Demattê, J.A.M., Morgan, C.L.S., 2015. Predicting clay content on field-moist intact tropical soils using a dried, ground VisNIR library with external parameter orthogonalization. Geoderma 259-260, 196-204.

Aenugu, H. P. R., Kumar, D. S., Srisudharson, N. P., Ghosh, S. S., Banji, D., 2011. Near infra red spectroscopy–an overview. International Journal of ChemTech Research, 3(2), 825-836.

Anderson, T.H., 2003. Microbial eco-physiological indicators to asses soil quality. Agr. Ecosyst. Environ. 98, 285–293.

Andrews, S.S., Karlen, D.L., Mitchell, J.P., 2002. A comparison of soil quality indexing methods for vegetable production systems in Northern California. Agriculture, Ecosystems Environment, 90(1), 25-45.

Arshad, M.A., Martin, S., 2002. Identifying critical limits for soil quality indicators in agro-ecosystems. Agr. Ecosyst. Environ. 88, 153–160

Bartlett, R. J., and James, B. R., 1993. Redox chemistry of soils. Adv. Agron. 50, 151–208.

Bartlett, R. J., and James, B. R., 1995. System for categorising soil redox status by chemical field testing. Geoderma 68, 211–218

Baumgardner, M. F., Silva, L. F., Biehl, L. L., Stoner, E. R., 1985. Reflectance properties of soils. Advances in Agronomy 38, 1–44.

Bédidi, A., Cervelle, B., Madeira, J., Pouget, M., 1992. Moisture effects on visible spectral characteristics of lateritic soils. Soil Science, No 153, pp. 129-141.

Bellinaso, H., Dematte, J.A.M., Araujo, S.R., 2010. Spectral library and its use in soil classification. Braz. J. Soil Sci. 34, 861-870.

Ben Dor, E., Patkin, K., Banin, A., Karnieli, A., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data. A case study over clayey soils in Israel. Int. J.Remote Sens. 23, 1043e1062

Bending, G.D., Turner, M.K., Rayns, F., Marx, M.C., Wood, M., 2004. Microbial and biochemical soil quality indicators and their potential for differentiating areas under contrasting agricultural management regimes. Soil Biol. Biochem. 36, 1785–1792.

Ben-Dor, E., Banin, A., 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. Soil Science Society of America Journal 59(2), 364-372.

Ben-Dor, E., Heller, D., Chudnovsky, A., 2008. A novel method of classifying soil profiles in the field using optical means. Soil Science Society of America Journal 72(4), 1113-1123.

Blanco, M., Villarroya, I., 2002. NIR spectroscopy: a rapid-response analytical tool. TrAC Trends in Analytical Chemistry, 21(4), 240-250.

Blum, W. E. H., 1993. Soil protection concept of the council of Europe and integrated soil research. In ‘‘Integrated Soil and Sediment Research: A Basis for Proper Protection’’ (H. J. P. Eijsackers and T. Hamers, Eds.), pp. 37–47. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Blum, W., Schad, P., Nortcliff, S., 2017. Essentials of Soil Science

Bouma, J., 1997. Soil environmental quality: A European perspective. J. Environ. Qual. 26, 26–31.

Bowers, S.A., Hanks R.J., 1965. Reflection of radiant energy from soils. Soil Science, 100, pp. 229-230.

Bowers, S.A., Smith S.J., 1965. Spectrophotometric determination of soil water content. Soil Sci. Soc. Amer. Proc., 36, pp. 978-980.

Brady, N., 1989. The nature and properties of soils. MacMillan Publishing Company, New York.

Brodsky, L., Klement, A., Penizek, V., Kodesova, R., Boruvka, L., 2011. Building soil spectral library of the Czech soils for quantitative digital soil mapping. Soil Water Res. 6, 165e172.

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

Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., De Deyn, G., de Goede, R., ... & Pulleman, M. (2018). Soil quality – A critical review. Soil Biology and Biochemistry, 120, 105-125.

Cañasveras Sanchez, J., Barron, V., Del Campillo, M., Viscarra Rossel, R.A., 2012. Reflectance spectroscopy: a tool for predicting soil properties related to the incidence of Fechlorosis. Span. J. Agric. Res. 10, 1133e1142

Carter, M.R., Gregorich, E.G., Anderson, D.W., Doran, J.W., Janzen, H.H., Pierce, F.J., 1997. Concepts of soil quality and their significance. in: Developments in Soil Science, (Eds.) E.G. Gregorich, M.R. Carter, Vol. Volume 25, Elsevier, pp. 1-19.

Cécillon, L., Barthès, B. G., Gomez, C., Ertlen, D., Génot, V., Hedde, M., ... & Brun, J. J., 2009a. Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). European Journal of Soil Science, 60(5), 770-784

Cècillon, L., Cassagne, N., Czarnes, S., Gros, R., Vennetier, M. Brun, J.J., 2009 b. Predicting soil quality indices with near infrared analysis in a wildfire chronosequence. Science of the Total Environment, 407, 1200–1205.

Celis-Ceuster, A.M., 1980. Ground truth radiometry (Exotech) on bare and overgrown Belgian soils. Pédologie, 30, pp. 43-63.

Cervelle, B., Flay, N., 1995. Diffuse reflectance properties of mineral bearing surfaces. Coloquio Internacional 'Propietades espectrales y teledeteccion de los suelos y rocas del visibile al infrarrojo medio', La Serena, 24-27 Abril 1995, pp. 1-26.

Chabrillat, S., Goetz, A. F., Krosley, L., Olsen, H. W. , 2002. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. Remote sensing of Environment, 82(2-3), 431-445.

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.

Cierniewski, J., 1985. Relation between soil moisture tension and spectral reflectance of different soils in visible and near-infrared range. In: 3ème Coll. Int. "Signatures spectrales d'objets en télédétection". Laes Arce, 16-20/12/85, pp. 429-432.

Clark, R. N., 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. Manual of remote sensing, 3(3-58), 2-2.

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. Journal of Geophysical Research 95, 12653–12680.

Cohen, M. J., Dabral, S., Graham, W. D., Prenger, J. P., Debusk, W. F..2006. Evaluating ecological condition using soil biogeochemical parameters and near infrared reflectance spectra. Environmental Monitoring and Assessment, 116(1-3), 427-457.

Curtaz F., and Zanini E.,. 2012. "Guida pratica di pedologia". Institut Agricole Régional: 5-125.

Dalal, R.C., Henry, R.J., 1986. Simultaneous Determination of Moisture, Organic Carbon, and Total Nitrogen by Near Infrared Reflectance Spectrophotometry. Soil Science Society of America Journal 50(1), 120-123.

Daniel, K.W., Tripathi, N.K., Honda, K., 2003. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). Australian Journal of Soil Research 41, 47–59.

Davidson, D.A. 2000. Soil quality assessment: recent advances and controversies. Progress in Environmental Science, 2, 342-350.

Dematte, J. A. M., Campos, R. C., Alves, M. C., Fiorio, P. R., Nanni, M. R., 2004. Visible-NIR reflectance: a new approach on soil evaluation. Geoderma 121 (1-2), 95–112.

Dematte, J.A.M., Terra, F.d.S., Quartaroli, C.F., 2012. Spectral behavior of some modal soil profiles from São Paulo State, Brazil. Bragantia 71, 413-423.

Dick, R.P., 1992. A review: long-term effects of agricultural systems on soil biochemical and microbial parameters. Agr. Ecosyst. Environ. 40, 25–36.

Dick, W. A., Thavamani, B., Conley, S., Blaisdell, R., Sengupta, A., 2013. Prediction of β-glucosidase and β-glucosaminidase activities, soil organic C, and amino sugar N in a diverse population of soils using near infrared reflectance spectroscopy. Soil Biology and Biochemistry, 56, 99-104.

Doran, J.W., Parkin, T.B. 1994. Defining and Assessing Soil Quality. in: Defining Soil Quality for a Sustainable Environment, (Eds.) J.W. Doran, D.C. Coleman, D.F. Bezdicek, B.A. Stewart, SSSA. Madison, WI, pp. 3-21.

Doran, J.W., Parkin, T.B., 1996. Quantitative Indicators of Soil Quality: A Minimum Data Set. in: Methods for Assessing Soil Quality, (Eds.) J.W. Doran, A.J. Jones, Soil Science Society of America, pp. 25-37.

Drury S.A.,1993. Image interpretation in geology. Chapman & Hall, London.

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

Dyrby, M., S.B. Engelsen, L. Nørgaard, M. Bruhn, and L. Lundsberg-Nielsen, 2002. "Chemometric quantitation of the active substance (containing C≡ N) in a pharmaceutical tablet using near-infrared (NIR) transmittance and NIR FT-Raman spectra." Applied Spectroscopy 56 (2002): 579-585. Print

Eastes J.W., 1989. Spectral properties of halite-rich minerals mixtures: Implications for middle infrared remote sensing of highly saline environments. Remote Sens. Environ., 27, pp. 289-304.

Fideñcio, P.H., Poppi, R.J., De Andrade, J.C., 2002. Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy. Analytica Chimica Acta 453, 125– 134.

Fiedler, S., Vepraskas, M. J., Richardson, J. L., 2007. Soil redox potential: importance, field measurements, and observations. Advances in Agronomy, 94, 1-54.

Flaig W. Beutelspacher H., Rietz E. (1975) – Chemical composition and physical properties of humic substances. In "Gieseking J.E. Ed.", Soil components vol. 1, Organic components, pp. 1-211.

Gaffey, S.J., 1986. Spectral reflectance of calcite minerals in the visible and near infrared (0.35-2.55 microns): calcite, aragonite, and dolomite. Am. Mineral. 71:151-162.

Gaffney, J.S., Marley, N.A., Clark, S.B., 1996. Humic and fulvic acids and organic colloidal materials in the environment. In: Gaffney, J.S., Marley, N.A., Clark S.B. (Eds.), Humic and Fulvic Acids.Isolation, Structure and Environmental Role. ACS SymposiumSeries 651.

Gao, S., Tanji, K. K., Scardaci, S. C., and Chow, A. T., 2002. Comparison of Redox indicators in a paddy soil during rice-growing season. Soil Sci. Soc. Am. J. 66, 805–817.

Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: a tutorial. Analytica Chimica Acta 185, 1– 17.

Giacometti, C., Demyan, M. S., Cavani, L., Marzadori, C., Ciavatta, C., Kandeler, E., 2013. Chemical and microbiological soil quality indicators and their potential to differentiate fertilization regimes in temperate agroecosystems. Applied Soil Ecology, 64, 32-48.

Gobrecht, A., Roger, J.M., Bellon-Maurel, V., 2014. Major issues of diffuse reflectance NIR spectroscopy in the specific context of soil carbon content estimation: a review. Adv. Agron. 123, 145e175.

Goetz, A.F., S. Chabrillat, and Z. Lu., 2001. Field reflectance spectroscopy for detection of swelling clays at construction sites. Field Anal. Chem. Technol. 5:143-155.

Gogé, F., Joffre, R., Jolivet, C., Ross, I., Ranjard, L., 2012. Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database. Chemom. Intell. Lab. 110, 168e176.

Guerrero, C., Zornoza, R., Gómez, I., Mataix-Beneyto, J., 2010. Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. Geoderma, 158(1-2), 66-77.

Koch, A., McBratney, A., Adams, M., Field, D., Hill, R., Crawford, J., ... & Angers, D. (2013). Soil security: solving the global soil crisis. Global Policy, 4(4), 434-441.

Harris, R. F., Karlen, D. L., and Mulla, D. J., 1996. A conceptual framework for assessment and management of soil quality and health. In ‘‘Methods for Assesssing Soil Quality’’ (A. J. Jones and J. W. Doran, Eds.), Vol. 49, pp. 61–82. SSSA, Madison, WI Special Publication.

Hartemink, A.E., Minasny, B., 2014. Towards digital soil morphometrics. Geoderma 230, 305-317

He, J., Song, H., García Pereira, A., Hernandez Gomez, A., 2005. A new approach to predict N, P, K and OM content in a loamy mixed soil by using near infrared reflectance spectroscopy. Adv. Intell. Comput. 3644, 859e867.

Hedde, M., Lavelle, P., Joffre, R., Jim´enez, J.J. Deca¨ens, T. 2005. Specific functional signature in soil macro-invertebrates biostructures. Functional Ecology, 19, 785–793

Hedley, C., Roudier, P., Maddi, L., 2014. VNIR Soil Spectroscopy for Field Soil Analysis. Communications in Soil Science and Plant Analysis 46(sup1), 104-121.

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, pp. 865-872.

Hunt G.R., Salisbury J.W., 1970. Visible and near infrared spectra of minerals and rocks. I. Silicate minerals. Modern Geology, 1, pp. 283-300

Hunt G.R., Salisbury J.W.,Lenhoff C.J., 1971. Visible and near infrared spectra of minerals and rocks. III. Oxides and hydroxides. Modern Geology, 2, pp. 195-205

Hunt, G.R., 1977. Spectral signatures of particulate minerals in the visible and near infrared. Geophysics 42, 501–513.

Irons J.R., Weismiller R.A., Petersen G.W. 1989. Soil reflectance. In ‘Theory and Applications of optical remote sensing’ (G. Asrar, Ed.), Wiley, New York, pp 66- 106.

Islam, K., McBratney, A., Singh, B., 2005. Rapid estimation of soil variability from the convex hull biplot area of topsoil ultra-violet, visible and near-infrared diffuse reflectance spectra. Geoderma 128 (3-4), 249–257.

Jackson M.L., Clayton R.N., Violante A. and Violante P., 1981. Eolian influence on Terra Rossa soils of Italy traced by quartz oxygen isotopic ratio. pp. 293-301. In:Int. Clay Conf. 1981 (Eds. H. van Olphen, and F. Veniale). Elsevier Sci. Pub. Co., Amterdam, The Netherland.

Jarvis, N. J., 2007. A review of non-equilibrium water flow and solute transport in soil macropores: Principles, controlling factors and consequences for water quality. Eur. J. Soil Sci. 58, 523–546.

Jenny, H.,1980. The Soil Resource Springer, Berlin German Federal Republic.

Karlen, D. L., Mausbach, M. J., Doran, J. W., Cline, R. G., Harris, R. F., and Schuman, G. E., 1997. Soil quality: A concept, definition, and Framework for Evaluation. Soil Sci. Soc. Am. J. 61, 4–10.

Karmonov, I.I., Rozhkov, V.A., 1972. Experimental determination of quantitative relationships between the color characyeristics of soils and soil constituents. Sov. Soil Sci. (Eng. Trans.), 4, pp. 666-674.

Knadel, M., Deng, F., Thomsen, A., Greve, H.M., 2012. Development of a Danish National Vis-NIR Soil Spectral Library for Soil Organic Carbon Determination. Digital Soil Assessments and beyond. CRC Press, Boca Raton, FL, pp. 403-408.

Körschens, M., 2006. The importance of long-term field experiments for soil science and environmental research – a review. Plant Soil Environ. 52, 1–8

Krishnan, P., Alexander, J.D., Butler, B.J., Hummel, J.W., 1980. Reflectance Technique for Predicting Soil Organic Matter. Soil Science Society of America Journal 44(6), 1282-1285.

Leone, A.P., 2000. Spettrometria e valutazione della riflettanza spettrale dei suoli nel dominio ottico 400-2500 nm. AITinforma- Riv. It. Telerilevamento, No. 19, pp. 3- 28.

Madeira Netto, J., Da, S., 1991. Etude quantitative des relations constituants minéralogiques – réflectance diffuse des lotosols brésiliens. Applications a l'utilization pédologique des donnès satellitares TM (région de Brasilia). Thèse de doctorat, Universié Pierre et Marie Curie, Paris, p.236.

Malley, D.F., L. Yesmin, and R.G. Eilers, 2002. Rapid analysis of hog manure and manure-amended soils using near-infrared spectroscopy. Soil Sci. Soc. Am. J. 66:1677-1686

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

McDowell, M.L., Bruland, G.L., Deenik, J.L., Grunwald, S., 2012. Effects of Subsetting by Carbon Content, Soil Order, and Spectral Classification on Prediction of Soil Total Carbon with Diffuse Reflectance Spectroscopy. Applied and Environmental Soil Science 2012, 14.

Minasny, B., McBratney, A. B., Bellon-Maurel, V., Roger, J. M., Gobrecht, A., Ferrand, L., Joalland, S., 2011. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. Geoderma, 167, 118-124.

Morgan, C.L., T.H. Waiser, D.J. Brown, and C.T. Hallmark. 2009. Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. Geoderma 151:249-256.

Morris, R. V., Lauer, H. V., Lawson, C. A., Gibson, E. K., Nace, G. A., Stewart, C., 1985. Spectral and other physicochemical properties of submicron powders of hematite (α-Fe2O3), maghemite (γ-Fe2O3), magnetite (Fe3O4), goethite (α-FeOOH), and lepidocrocite (γ-FeOOH). Journal of Geophysical Research: Solid Earth, 90(B4), 3126-3144.

Mouazen, A. M., Karoui, R., De Baerdemaeker, J., Ramon, H., 2005. Classification of soil texture classes by using soil visual near infrared spectroscopy and factorial discriminant analysis techniques. Journal of near infrared spectroscopy, 13(4), 231-240.

Mouazen, A. M., Karoui, R., De Baerdemaeker, J., Ramon, H., 2006. Classification of soils into different moisture content levels based on VIS-NIR spectra. In 2006 ASAE Annual Meeting (p. 1). American Society of Agricultural and Biological Engineers.

Mougenot, B., Pouget M., Epena G.F., 1993. Remote sensing of salt affected soils. Remote Sensing Reviews. No 7, pp. 241-259.

Mulders, M.A., 1987. Remote sensing in soil science. Developments in soil science, Vol. 15 Elsevier, Amsterdam, p. 379.

Myers, V.I., Allen ,W.A., 1968. Electro-optical remote sensing methods as non-destructive testing and measuring techniques in qgriculture. Appl. Opt., 7, p p. 1818-1838.

Næs, T., Isaksson, T., Fearn, T., Davies, T., 2002. A user friendly guide to multivariate calibration and classification. NIR publications.

Næs, T., Isaksson, TKowalski B.R., 1990. Locally weighted regression and scater-correction for Near.Infrared Reflectance Data, Anal.Chem. 62, 664-673.

Nguyen, T.T., Janik, L.J., Raupach, M., 1991. Diffuse Reflectance Infrared Fourier Transform (DRIFT) spectroscopy in soil studies. Australian Journal of Soil Research 29, 49–67.

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.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., ... & Dardenne, P. (2015). Soil spectroscopy: An alternative to wet chemistry for soil monitoring. In Advances in agronomy (Vol. 132, pp. 139-159). Academic Press.

O'Rourke, S.M., Holden, N.M., 2011. Optical sensing and chemometric analysis of soil organic carbon - a cost effective alternative to conventional laboratory methods? Soil Use Manage. 27, 143e155.

Obukhov, A.I., Orlov, D.S., 1964. Spectral reflectivity of the major soil groups and possibility of using diffuse reflection in soil investigations. Soviet oil Science, 1, pp. 174-184.

Odlare, M., Svensson, K., Pell, M., 2005. Near infrared reflectance spectroscopy for assessment of spatial soil variation in an agricultural field. Geoderma 126 (3-4), 193–202

Owen, T., 2000. Fundamentals of Modern UV-Visible Spectroscopy: A Primer. edit. Agilent Technologie, German.

Pajares, S., Gallardo, J.F., Masciandaro, G., Ceccanti, B., Marinari, S., Etchevers, J.D., 2009. Biochemical indicators of carbon dynamics in an Acrisol cultivated under different management practices in the central Mexican highlands. Soil Till. Res.105, 156–163

Palmborg, C. and A. Nordgren. 1996. Partitioning the variation of microbial measurements in forest soils into heavy metal and substrate quality dependent parts by use of near infrared spectroscopy and multivariate statistics. Soil Biol. Biochem. 28(6):711-720

Phogat, V.K., TomarRita, V.S., Dahiya R., 2015. Soil Physical Properties. In: Soil Science: An IntroductionEdition: FirstPublisher: Indian Society of Soil ScienceEditors: Rattan R.K., Katyal J.C., Dwivedi B.S., Sarkar A.K., Bhattachatyya Tapan, Tarafdar J.C., Kukal S.S.

Rao, M. A., Scelza, R., Gianfreda, L. (2014). Soil enzymes. Enzymes in agricultural sciences OMICS eBooks Group, 10-43.

Rasmussen, P.E., Collins, H.P., 1991. Long-term impacts of tillage, fertilizer, and crop residue on soil organic matter in temperate semiarid regions. Adv. Agron. 45, 93–134.

Reeves, D.W., 1997. The role of organic matter in maintaining soil quality in continuous cropping systems. Soil Till. Res. 43, 131–167.

Rossel, R. V., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., ... & Aïchi, H., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews, 155, 198-230.

Rowell, D. L., 2014. Soil science: Methods & applications. Routledge

Schloter, M., Dilly, O., Munch, J.C., 2003. Indicators for evaluating soil quality. Agr. Ecosyst. Environ. 98, 255–262.

Schwartz, G., Ben-Dor, E., Eshel, G., 2012. Quantitative analysis of total petroleum hydrocarbons in soils: comparison between reflectance spectroscopy and solvent extraction by 3 certified laboratories. Appl. Environ. Soil Sci. 11.

Shepherd, K.D., Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66, 988– 998.

Shibusawa, S., Imade Anom, S.W., Sato, S., Sasao, A., Hirako, S., 2001. Soil mapping using the real-time soil spectrophotometer. In: Grenier, G., Blackmore, S. (Eds.), ECPA 2001, Third European Conference on Precision Agriculture, vol. 1. Agro Montpellier, pp. 497– 508.

Shields, J.A., Paul, E.A., St Arnaud, R.J., Head, W.K., 1968. Spectrophotometric measurements of soil color and its relation to moisture and organic matter. Can. J. of Soil Sci., 46, pp. 271-280

Siesler, H.W., 2008. Basic principles of near-infrared spectroscopy. In: Burns, D. A., Ciurczak, E. W. (eds.), Handbook of Near-Infrared Analysis. CRC Press, Boca Raton, Fla.

Soil Survey Staff, 1999. Soil Taxonomy. Agriculture Handbook 436.

Soil Survey Staff, 2006. Keys to soil taxonomy. USDA, NATURAL Resource Conservation Services, Washington, USA, pp. 332

Sorensen, L.K., Dalsgaard, S., 2005. Determination of clay and other soil properties by near infrared spectroscopy. Soil Sci. Soc. Am. J. 69, 159e167.

Springsteen A., 1994. A Guide to Reflectance Spectroscopy. Labsphere Tech Guide, North Sutton, NH

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. doi:10.1016/j.geoderma.2010.04.008.

Stenberg, B., Jonsson, A., and Bo¨ rjesson, T. (2002). Near infrared technology for soil analysis with implications for precision agriculture. In ‘‘Near Infrared Spectroscopy: Proceedings of the 10th International Conference’’ (A. Davies and R. Cho, Eds.), pp. 279–284. NIR Publications, Chichester, UK.

Stenberg, B., Viscarra Rossel, R. A., Mouazen, A. M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. Advances in Agronomy 107, 163–215.

Stevenson, F.J. and Cole, M.A., 1999. Cycles of soil carbon, nitrogen, phosphorus, sulfur,micronutrients. In: Cycles ofSoil. Carbon, Nitrogen, Phosphorus, Sulfur,Micronutrients, (Eds., Stevenson, F.J. and Cole, M.A.). pp. 87-89. John Wiley andSons, New York.

Stoner E., Baumgardner M. F., 1981 Characteristics variations in reflectance of surface soils. Soil Sci. Soc. Am. J. 45, pp. 1161-1165.

Sudduth, K. A., Kitchen, N. R., Kremer, R. J., 2009. VNIR spectroscopy estimation of soil quality indicators. In 2009 Reno, Nevada, June 21-June 24, 2009 (p. 1). American Society of Agricultural and Biological Engineers.

Todorova, M., Atanassova, S., Lange, H., Pavlov, D., 2011. Estimation of total N, total P, pH and electrical conductivity in soil by near-infrared reflectance spectroscopy. Agr. Sci. Tech. 3, 50e54.

Toth, G., Jones, A., Montanarella, L., 2013. The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. Environ. Monit. Assess. 185, 7409e7425.

van der Meer F., 1995. Imaging spectrometry and Ronda peridotites. PhD Thesis, University of Wageningen, NL.

Vasques, G.M., Demattê, J.A.M., Ramírez-López, L., Terra, F.S., 2014. Soil classification using visible/near-infrared diffuse reflectance spectra from multiple depths. Geoderma 223-225, 73-78.

Vasques, G.M., Grunwald, S., Harris, W.G., 2010. Spectroscopic models of soil organic carbon in Florida, USA. J. Environ. Qual. 39, 923-934.

Velasquez, E., Lavelle, P., Barrios, E., Joffre, R. Reversat, F. 2005. Evaluating soil quality in tropical agroecosystems of Colombia using NIRS. Soil Biology & Biochemistry, 37, 889–898.

Vepraskas, M. J., Richardson, J. L., Tandarich, J. P., Teets, S. J., 1999. Dynamics of hydric soil formation across the edge of a created deep marsh. Wetlands 19, 78–89.

Vergnoux, A., Guiliano, M., Le Dréau, Y., Kister, J., Dupuy, N., Doumenq, P., 2009. Monitoring of the evolution of an industrial compost and prediction of some compost properties by NIR spectroscopy. Science of the Total Environment, 407(7), 2390-2403.

Vinogradov, B.V. 1981. Remote sensing of the humus content of soils. Sov. Soil Sci. (Engl. Transl.), 6, pp. 103- 113.

Viscarra Rossel, R.A., Walvoort, D.J.J., 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, 59e75.

Viscarra-Rossel, R.A., S.R. Cattle, A. Ortega, and Y. Foua. 2009. In situ measurements of soil colour, mineral composition and clay content by Vis-NIR spectroscopy. Geoderma 150:253-266.

Waiser, T., C.L.S. Morgan, D.J. Brown, and C.T. Hallmark. 2007. In situ characterization of soil clay content with VNIR diffuse reflectance spectroscopy. Soil Sci. Soc. Am. J. 71:389-396.

Winding, A., Hund-Rinke, K., Rutgers, M., 2005. The use of microorganism in ecological soil classification and assessment concepts. Ecotox. Environ. Saf. 62, 230–248

Zanella, A., Tomasi, M., De Siena, C., Frizzera, L., Jabiol, B., Nicolini, G., ... & Pizzeghello, D., 2001. Humus forestali. Manuale di ecologia per il riconoscimento e l'interpretazione. Applicazione alle faggete. Centro di Ecologia Alpina, Trento, 321.

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K. M., Arcenegui, V., Mataix-Beneyto, J., 2009. Changes in soil microbial community structure following the abandonment of agricultural terraces in mountainous areas of Eastern Spain. Applied Soil Ecology, 42(3), 315-323..

CAP.2

ESTIMATION OF ANDIC PROPERTIES FROM VISNIR DIFFUSE REFLECTANCE SPECTROSCOPY IN EUROPEAN VOLCANIC SOILS

2.1. Introduction

Soils formed from volcanic materials are generally associated with the occurrence of volcanic effusions and their estimate extension is around 124 million hectares, or 0.84% of the world land surface (IUSS Working Group, 2014). However, the surface covered by volcanic soils, or by soil that has one or more andic properties, may be more extensive than currently estimated. In fact, the evidence of volcanic origin is sometimes difficult to verify, since andic properties can rapidly be modified by different pedogenetic processes (Wada, 1985; Parfitt and Kimble, 1989; Ugolini and Dahlgren, 2002). In Italy, for instance, the distribution of soils formed on volcanic parent materials was found by some authors to be more widespread than expected from current soil maps (Buurman et al., 2003; Iamarino and Terribile, 2008; Colombo et al., 2014).

Volcanic soils have a high capacity to retain water, nutritional elements (cations and anions) and organic compounds, making volcanic soils important for their high fertility (Takahashi and Shoji, 2002). Volcanic soils, for their peculiar properties, have been recognized as a major pedological unit and named "Andosols" (IUSS Working Group, 2014). They are characterized by the abundance of poorly crystalline materials associated with high soil organic matter (SOM), low bulk density, high phosphorus retention, and large water retention capacity (Wada and Aomine, 1973). Andosols generally show dark organic upper horizons and deep lighter colored horizons, related to the chemical composition of the pyroclastic deposit (Shoji et al., 1993). The carbon storage capacity of andic soils is closely related to the high surface areas of poorly crystalline constituents (i.e. pyrophosphate and oxalate extractable Al and Fe) that are available for the sorption of organic matter (Matus, et al., 2014).

According to the World Reference Base for soil resources (WRB, 2014), the following parameters are diagnostic to detect andic and vitric properties, on which basis volcanic soils are classified (Dahlgren et al., 2004): acid oxalate extractable aluminum (Al_o) and iron (Fe_o) content, combined as $Al_o+1/2Fe_o$ (>2% for andic properties, >0.4% for vitric properties); Silicon (Si_o) used to discriminate between Silandic ($Si_o \geq 0.6\%$) and Aluandic Andosols; phosphate retention (PR>70%); bulk density (>0.9 kg dm^{-3}); Aluminum and Iron content determined through pyrophosphate and dithionite extraction (Al_d and Fe_d; Al_p and Fe_p), in particular, Al_d/Al_o is used to discriminate Silandic (<0.5) from Aluandic soils (García-Rodeja et al., 2004), Al_o - Al_p represents the mineral reactive Al fraction. Other relevant parameters are: pH in NaF (pHNaF>9.5) for a preliminary field identification of presence of allophanic products or/and Al-organic matter complexes (IUSS Working Group, 2014); base saturation (BS) and cation exchangeable capacity (CEC) (Madeira et al., 2003); clay and total organic carbon (TOC) content (Shoji et al., 1985); color (Brown et al., 2006).

An alternative and fast method to measure soil characteristics is the use of visible and near-infrared reflectance (VisNIR) spectroscopy. In this region, the radiation is absorbed by the different chemical bonds of the compounds present in the sample. Moreover, the radiation is absorbed in accordance with the concentration of these compounds. Thus, VisNIR spectra contains information about the compositional qualities of a soil sample and researchers are increasingly turning to reflectance spectroscopy to overcome the expense of traditional soil survey and analysis, and the limitations of pedotransfer functions (Bédidi et al., 1992; Ben-Dor et al., 2008; Stevens et al., 2013; Wang et al., 2015). VisNIR spectra can provide useful indicators for mapping, classifying and monitoring soils for a large range of properties (Ben-Dor and Banin, 1995; Viscarra Rossel et al., 2009; Priori et al., 2016) with an improvement in the speed and resolution of soil physical and chemical analysis (Nanni and Demattê, 2006; Cañasveras et al., 2010; Ge et al., 2011; Nocita et al., 2015). Spectral analysis of soil cores with field or laboratory spectrometers could provide a new tool of automated, rapid and objective profile evaluation, similarly to the approach that is now being developed for mineral characterization of geological cores (Calvin et al., 2015). In addition, new imaging spectrometers offer the prospect of detailed raster-based mapping of surface soil properties with higher spatial resolution than is possible with the current approaches (Kodaira and Shibusawa, 2013).

Regression methods such as multiple linear (MLR), principal component (PCR), and partial least squares (PLSR) can be used to predict soil attributes through reflectance spectroscopy (Stone and Brooks, 1990; Minasny, and McBratney, 2008). Among which, PLSR should be highlighted as an efficient algorithm for predictive models of various chemical (Viscarra Rossel et al., 2009), physical (Gomez et al., 2013), and biological (Cécillon et al., 2009) soil attributes. Some studies investigated the VisNIR spectra with a data mining approach. Some of these methods outperformed PLSR, for instance, back propagation neural network (Mouazen et al. 2010) and boosted regression trees (BRT) (Brown et al. 2005). According to Viscarra Rossel and Behrens (2010) predictions of soil organic matter (SOM), clay content and pH was better with the supported vector machine (SVM) than with other multivariate data mining algorithms, while PLSR and multivariate adaptive regression splines (MARS) produced competitive predictions compared with SVM. On the other hand, McDowell et al. (2012a) did not find a significant difference among PLSR and random forest (RF) method, and Grunwald and Xiong (2014) pointed to PLSR being a better estimator than SVM and RF for SOM content.

Currently, only few studies investigated andic soils with diffuse reflectance spectroscopy (DRS) in the VisNIR. Good results were obtained by Bellino et al. (2015), who found that few observations of andic soil profiles (20 samples from 3 stations) were enough to predict Fe and Al extracted by sodium pyrophosphate (RPD>3.5) and, to a lesser extent, Fe and Al extracted by ammonium oxalate. On the other hand, the presence of Andosols seems to negatively affect TOC predictability (Kühnel and Bogner, 2017; McDowell et al. 2012b), which is generally high (Sellitto et al., 2007, Zhao et al., 2012, Bellino et al., 2015, Kinoshita et al., 2016; Bonett et al., 2016).

The general objective of this study was to investigate the usefulness of DRS for the classification of soils belonging to a European set of profiles on volcanic rocks. The first activity of this research was to investigate the efficiency of DRS in classifying soil samples according to their different andic degrees. Then the research focused on predictive models to quantitatively estimate the andic properties by DRS. Since the lack of studies about algorithms to predict andic properties by DRS, in this study we compared two different methods: partial least square regression (PLSR) and supporter vector machine (SVM). PLSR is the most common algorithm to calibrate VisNIR spectra to soil properties (Wold et al., 1983); while SVM is recognized to be able to manage also an eventual non-linear input-output relationship and may outperform several machine learning algorithms in the prediction of some soil properties (Viscarra Rossel and Behrens, 2010).

2.2. Materials and Methods

2.2.1 Soil selection and analysis

We examined 67 soil horizons, sampled from 14 profiles located in different European countries (Fig. 2.1). The soils were the reference pedons of the European funded scientific action "COST-622: Soil Resources of European Volcanic Systems" (Arnalds et al., 2007; Table 2.1). The soil descriptions are available at http://www.rala.is/andosol/profiles/. The studied soils were developed on volcanic materials from Italy (profiles EUR01 to 04), Azores Islands, Portugal (EUR05, EUR06), Iceland (EUR07 to 09), Canary Islands, Spain (EUR10), France (EUR16), and Hungary (EUR18 to 19) (Table 1). Fe, Al, and Si were extracted by sodium dithionite citrate (Fed, Ald), ammonium oxalate (Feo, Alo, Sio) and pyrophosphate sodium (Fep, Alp). The suspension was centrifuged at 2500 rpm for 15 minutes and the supernatant was filtered before AAS determination (Buurman et al., 1996). Total organic carbon (TOC) content was determined using a LECO CHN 1000 Analyzer. Soil chemical extraction and routine soil data came from the COST soil chemical data base of the reference pedons (García-Rodeja et al., 2007).

2.2.2. Sample preparation and DRS measurement

The samples were air-dried, sieved at < 2 mm, and vigorously ground in an agate mortar for at least 10 minutes, up to a size fraction smaller than 1 μm, to exclude the influence of micro-aggregation. The samples were then gently pressed in a circular hole (diameter of 10 mm) against unglazed white paper, to avoid undesired particle orientation. DRS measurements were obtained by using a Jasco 560 UV-visible spectrophotometer, equipped with an integrating sphere, following the method proposed by Torrent and Barrón (1993). Each spectrum was made up of 1151 wavelengths, from 200 to 2500 nm. VisNIR spectra were recorded as percent reflectance (R%). Data acquisition was carried out by means of the Jasco software (Model VWTS-581 version 2.00A). The standard white was obtained by barium sulfate [Merck DIN 5033].

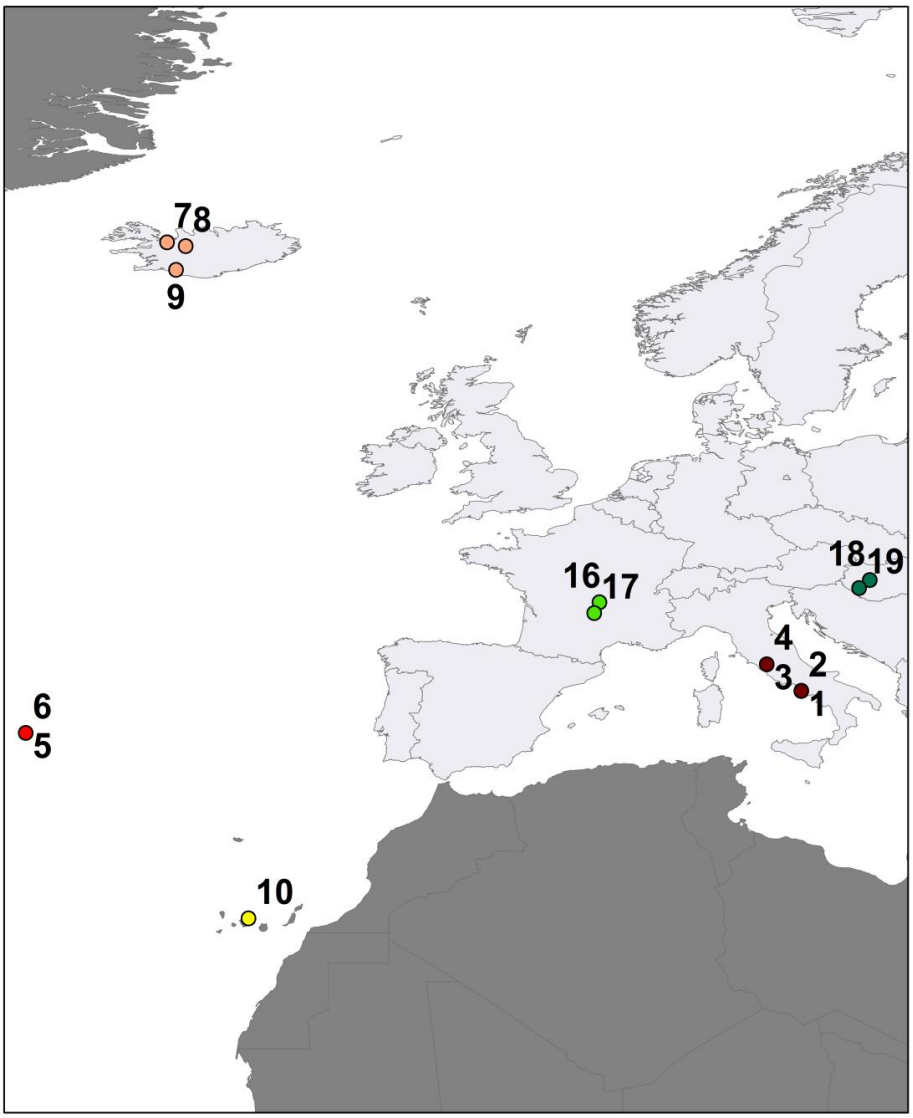


Figure 2.1. Geographic distribution of European volcanic soils. 1-4: Italy; 5,6: Portugal (Azores); 7,8,9: Iceland; 10: Spain (Tenerife); 15,16: France; 18,19: Hungary.

Table 2.1. Soil characteristics of the 14 reference pedons of the European "COST-622" action sampled in different European countries.

Profiles	Location	Soil Depth (cm)	Altitude	Climate: (USDA, 2006)	Parent material	Soil Classification WRB (WRB-FAO, 2014)
EUR01	Italy, Gauro	120	103	thermic xeric	Pyroclastic trachytic ash deposits	Humi-Tephric Regosol (Eutric)
EUR02	Italy, Gauro	92	225	mesic xeric	Pyroclastic trachytic deposits	Eutri-Humic Cambisol
EUR03	Italy, Vico	125	700	mesic udic	Lava, tephritic-phonolitic with leucite	Fulvi-Silandic Andosol (Dystric)
EUR04	Italy, Vico	105	722	mesic udic	Lava, tephritic-phonolitic with leucite	Fulvi-Silandic Andosol (Dystric)
EUR05	Portugal, Azores-Faial	145	510	mesci udic	Pyroclastic material	Hyperdystri-Silandic Andosol
EUR06	Portugal, Azores- Pico	140	400	mesci udic	Basaltic pyroclastic material	Hydri-silandic Andosol (Umbric and Acroxic)
EUR07	Iceland, Route 1	100	40	Cryic/Frigid udic	Eolian and basaltic tephra sediments	Orthidystri-Vitric Andosol
EUR08	Iceland, Audkuluheidi	65	400	Cryic udic	aeolian and volcanic ash	Dystri-Vitric Andosol
EUR09	Hella	230	50	Cryic/ Frigid udic	aeolian, tephra and organic materials	Umbri-Vitric Andosol (Pachic and Orthidystric)
EUR10	Spain, Tenerife-Tacoronte	220	1130	Mesic udic	Basaltic ashes	Umbri-Silandic Andosol (Hyperdystric)
EUR16	France, Puy de La Vache.	120	1000	mesic udic	Volcanic scoria	Umbri-Silandic Andosol (Endoskeletic and Endoeutric)
EUR17	France, Buron du Perle	90	1080	mesic udic	Colluvium of trachyandesitic rock	Aluandi-Silandic Andosol (Umbric and Acrudoxic)
EUR18	Hungary, Tihanny Peninsula	70	180	mesic xeric	Basaltic tuff	Pachi-Endotephric Phaeozem (Siltic)
EUR19	Hungary, Badacsony	50	420	mesic xeric	Pyroclastic deposits; basaltic tuff	Skeletic Umbrisol

2.2.3. Preprocessing analysis

Descriptive statistics were computed to synthesize the main features of the studied profiles. Chemometric calibrations and validations were performed after converting reflectance into absorbance measurements through [log10 (1/Reflectance)]. We tested several spectroscopic preprocessing algorithms, namely multiplicative scatter correction (MSC) and the standard normal variate transformation (SNV) to reduce noise or enhance the spectra (Geladi et al., 1985; Barnes et al., 1989), however, since they did not improve the modelling, we ultimately did not use them.

The VisNIR reflectance spectra of the soil samples (n=67) are shown in Figure 2.2A. Figure 2.2B reports the reflectance rough data (R%) transformed into absorbance measurements. Subsequently, average centralization and Savinsky Golay second-order polynomial filter (Savitzky and Golay, 1964) were applied as pre-treatments for the VisNIR range, for the correction of light scattering variations of the reflectance spectra.

2.2.4 Discriminating soils with different andic properties

The taxonomic classification of profiles together with the analytical data of their horizons were used for grouping the samples. Thus, the groups included samples belonging to profiles with similar level of expression of andic properties.

Supervised chemometric classification was carried out by a discriminant analysis (DA) performed for NIR data over the mentioned groups. The discriminant functions produced for a given sample the probability of membership of each respective andic group. A given sample was assigned to the andic group with the highest probability. The main purpose was to confirm the sensitivity of soil spectral characteristics against the soil andicity, so that soil samples could be correctly classified by NIR as belonging to a specific andic group.

For the compression of NIR data, a first principal component analysis (PCA) was applied to the NIR spectral matrix, extracting the first 40 principal components (PCs). An amount of 7 PCs were selected for the DA, in order to maintain a proportion between variables and samples (64 samples) and avoid the risk of overfitting. The selected PCs were the most correlated with the analytical andic data. The factor scores of the selected PCs were used as inputs for the discriminant analysis. The PCA was performed by the Unscrambler software (Camo, Inc.) and the DA by Statistica 10 software (StatSoft Inc, Tulsa, OK, USA).

2.2.5 Predictive models

Predictive regression models were performed with PLSR algorithms by the Unscrambler software (Camo, Inc.). Differently from PCA, the PLSR components are built considering both the response variables (RS) and the predictor variables (PV). The creation of the components and the regression takes place at the same step, to maximize the covariance between predictor (namely spectral information) and response variables.

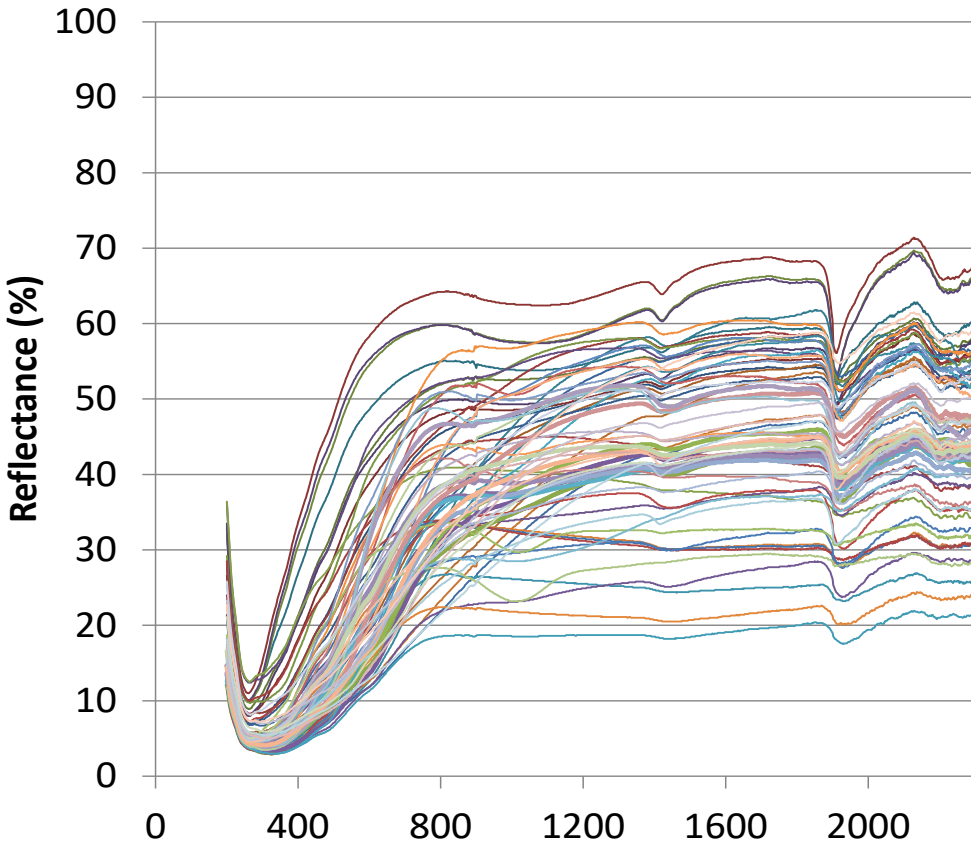
SVM was performed using Statistica 10 software (StatSoft Inc, Tulsa, OK, USA). SVM is a kernel-based learning method from statistical learning theory originally proposed by Vapnik (2000). To address non-linearity in input-output data relationships, SVM employs a kernel function to project the input data into a high dimensional space before performing the regression. In this way it is possible to derive a linear hyperplane as a decision function also for non-linear problems (Viscarra Rossel and Behrens, 2010). In SVM modeling, the Kernel radial basis function (Karatzoglou et al., 2008) was used, which allows learning of non-linear decision functions (Jain et al., 2012). The principal components obtained by PCA were used as predictive variables to overcome the problem of small sample sizes with many variables in assessing prediction error. We chose to use a number of PCs not higher than a tenth of the number of samples, for avoiding the risk of overfitting. Each model was built on the first seven PCs that were more correlated with the investigated soil parameter. The PCs were chosen separately by a Spearman test for each investigated soil parameter. The PCs with a correlation test less significant than p=0.15 were not included in the models.

A leave-one profile-out cross-validation (LOOCV) was performed to validate PLSR and SVM models. The dataset was split into a n number of folds, with n=total number of profiles, and n - 1 folds were used as training set and the last one profile as validation set. The operation was repeated until every profile was considered once in the validation set (Bellino et al., 2015). The LOOCV was used to overcome the problem of higher similarity between samples coming from the same profile, avoiding an optimistic estimation of the results.

2.2.6. Models comparison

The best calibrated PLSR and SVM models were compared for each estimated variable on the cross-validation dataset. We adopted the coefficient of determination (R^2), root mean square error of prediction (RMSEP), standard error of prediction (SEP), the BIAS and ratio of performance to deviance (Williams, 1987) (RPD), to evaluate the prediction performance of the calibrations. Although some criticism pointed out by some author about RPD when compared with SEP and BIAS (Bellon-Maurel, 2010), the RPD was chosen in this study for its widespread employment and so the possibility to compare the results with others. According to several authors, RPD values below 1.5 indicate a poor predictive model; in the range 1.5 - 2.0 they point to acceptable results, which however 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 (Chang et al., 2001; Dunn et al., 2002; Cozzolino and Moron, 2003).

a)



b)

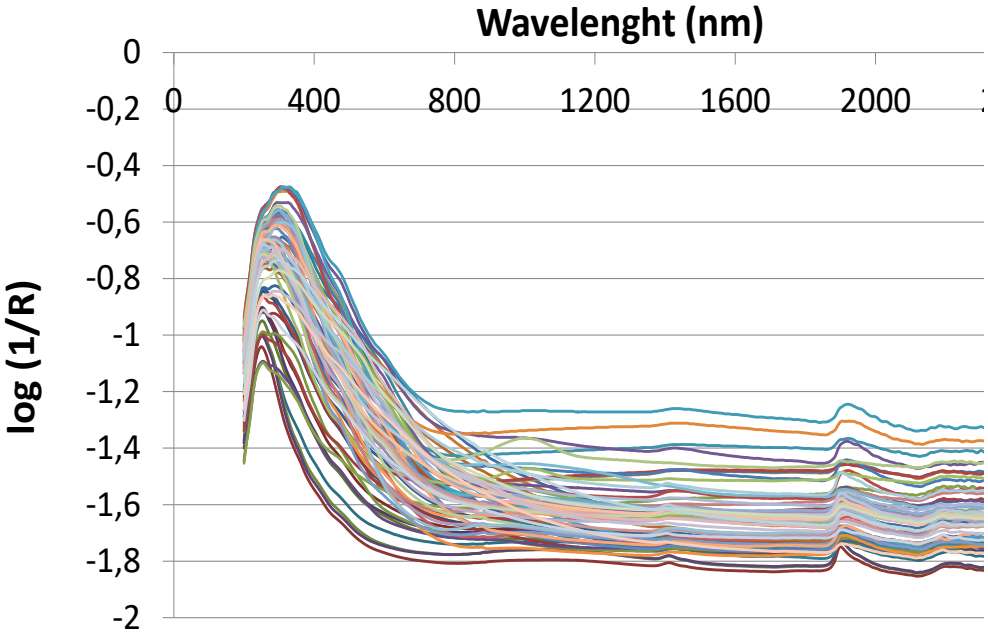


Figure 2.2. Summary of the 67 soil spectra in reflectance (a) and their transformation (b) into log10(1/Reflectance) units used in the modelling.

2.3. Results

2.3.1. Soil properties

The dataset was a collection of soil horizons of 14 soil profiles developed on different volcanic parent material in Europe. They were characterized by different degrees of andicity, from fitting the requirements for Tephric, Vitric, Allophanic, Aluandic and Silandic soil classification, to very low developed andic properties (Figure 1; Table 1).

In Mediterranean countries such as Spain, Italy and Portugal, there are a large variety of well-developed volcanic soils formed both from recent pyroclastic materials and volcanic lava. In Italy, two of the selected soil profiles showed very early pedogenesis and low andic properties (EUR01 and EUR02), both formed on rather fresh pyroclastic deposits. The third and fourth profiles, instead, were typical Andosols developed on leucites volcano lava rocks (EUR3 and EUR4). Prominent andic properties were especially observed in the profiles sampled in Azores-Pico in Portugal (EUR05 and EUR06) and Canaria Island in Spain, Tenerife (Silandic Andosols, EUR10). In north Europe many volcanic soils exhibited well expressed andic features, often with a peaty texture (wetland soils) formed on old tephra deposits. These Andosols were dominated by metal-humus complexes and poorly ordered Fe minerals and showed cryoturbation and hydromorphic features (EUR07, EUR08, EUR09, Iceland). Soil from France (EUR15 and EUR16) belonged to well-developed silandic Andodols on volcano lava rocks, while very low developed andic properties characterized soils from Hungary ranging from Pachi-Endotephric Phaeozem (EUR18) and skeletic Humbrisols (EUR19).

Table 2.2 shows the descriptive statistics of chemical properties and andic features of the 14 selected soil profiles and their 67 horizons. The TOC content for the whole dataset varied from 0.1 to 33.86 % with an average value of 7.46 %. Values for soil pH in KCl ranged from 4.00 to 6.10, on average 5.17, indicating a slightly acidic soil reaction. The percentage of clay in the samples ranged from 3.6% to 39% and averaged 17%, corresponding to the textural classes loamy sand, sandy loam, loam, silty loam. In the Regosols (EUR01), Cambisols (EUR02), Phaeozem (EUR18) and Humbrisols (EUR19) horizons, Alo and Feo values were low and generally with low variations along the profile. The andic index (Alo+1/2 Feo, %) always met the requirements set by the WRB (IUSS Working Group WRB, 2014), apart from EUR1, EUR2 and EUR18 and EUR19. Munsell hue ranged from 7.5YR (brown, brownish gray, orange) to 10YR (grayish yellow brown), Munsell values from 3.7 to 7.3, and chroma from 1.4 to 5.4. The color of most soils was highly sensitive to SOM and Al-Fe humic complexes resulted to increase in dark topsoils (3.7 value and chroma 1.4), but also in the B and C horizons with yellow orange colors 10YR 7/6. More developed soils, such as Silandic Andosols (EUR10, EUR16 and EUR17) generally occurred on older pyroclastic sediments.

Table 2.2. Statistical descriptions of the main chemical properties of the 67 horizons (14 soil profiles): total organic carbon (TOC); aluminum (Al_d), iron (Fe_d), manganese (Mn_d) extractable with dithionite citrate-bicarbonate; aluminum (Al_o), iron (Fe_o) and silicon (Si_o) extractable with ammonium oxalate; aluminum (Al_p), iron (Fe_p) extractable with pyrophosphate. C.V.: Coefficient of variation; St.D. standard deviation.

Parameter	**Min**	**Max**	**Mean**	**Variance**	**St.D.**	**Median**	**CV**
TOC (%)	0,10	33,86	7,46	52,20	7,23	5,84	0,969
Clay (%)	3,60	39,00	17,40	65,66	8,10	16,30	0,466
pH-NaF	7,60	11,50	9,47	1,73	1,32	9,60	0,139
pH-KCl	4,00	6,10	5,17	0,24	0,49	5,20	0,095
PR (%)	3	100	72	1001	32	88	0,437
CEC (cmol(+) kg^{-1})	11,10	94,10	41,12	333	18,24	38,20	0,444
BS (%)	1,90	100	35,68	1053	32,45	30,20	0,909
Al_d (%)	0,02	5,37	1,14	1,69	1,30	0,57	1,144
Fe_d (%)	0,14	13,07	2,94	9,49	3,08	1,60	1,048
Al_o (%)	0,09	10,07	2,81	5,81	2,41	2,38	0,859
Fe_o (%)	0,04	6,12	1,27	1,40	1,18	0,76	0,935
Si_o (%)	0,05	3,39	1,01	0,65	0,80	0,98	0,795
Al_o/Si_o	1,40	7,39	3,11	1,75	1,32	2,83	0,426
Al_p (%)	0,01	3,47	0,61	0,52	0,72	0,39	1,174
Fe_p (%)	0,01	4,24	0,50	0,77	0,88	0,17	1,744
$Al_o+1/2Fe_o$ (%)	0,11	11,70	3,44	7,98	2,82	2,85	0,821
$(Al_o-Al_p)/Si_o$	1,30	4,33	2,25	0,53	0,73	2,00	0,323
Allophane (%)	0,32	64,55	9,59	110	10,47	6,99	1,091

2.3.2. Spectra properties of the dataset

Soil VisNIR spectra covered a range from 35 to 65% of reflectance and the wide extension of this range points to a large difference in soil components (Fig. 2.3). Pronounced peaks were found around 1420, 1950, and 2200 nm. Figure 3 highlights a remarkable difference in intensity of reflectance both between and within soil profiles. Higher reflectance was observed in the horizons with lower TOC and B horizons had much stronger signals than O, A, or C

horizons. The horizons belonging to the same profile showed many more differences in the Vis range than in the NIR. The differences were increasing with soil depth and the absorption features changed between horizons at 1400, 1900 and 2200 nm. These absorption features are caused by the changes in clay mineralogy, which improves the stretching of the O–H bonds at 1400 nm, and by the Al–OH lattice structure at 2200 nm, due to the combination of the Al–OH bond plus O–H stretch of phyllosilicate minerals (Clark, 1999). Water molecules increase the NIR bands with a specific absorption peak at 1900, which is due to the combinations of the H–O–H bond with the O–H stretches, which is more pronounced in Halloysite minerals and in water rich minerals as Allophane. Allophane spectra exhibited a doublet near 1380 and 1400 nm due to O–H stretching, a broad band near 1920 nm due to H–O–H bond reflections (Bishop et al., 2013). Other differences are evident in the Vis region, with the changing in slope of the spectra curves, but are likely to be the result of other factors, such as texture, TOC and iron oxides.

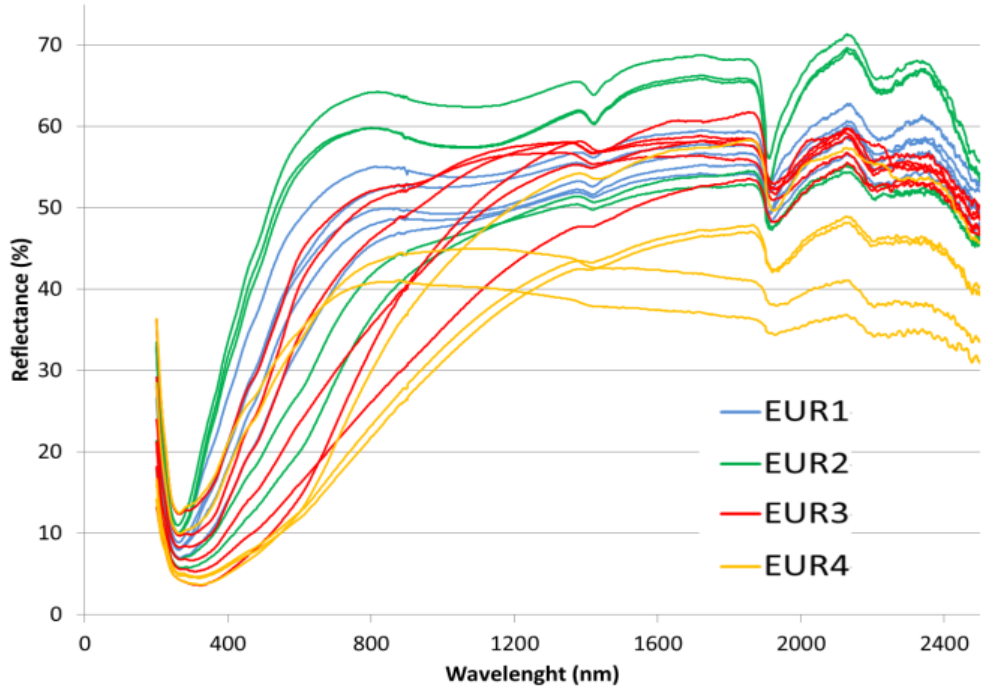


Figure 3. Spectra reflectance of the Italian profiles (EUR1-4).

2.3.3. PCA and discriminant analysis

The first four principal components accounted for 99.4% of the variation in the data. PC1 and PC2 accounted for 93.1% of the observed variance, whereas the explained variation of PC3 and PC4 was about 6.2%. Figure 2.4 shows the loadings weights of the wavelengths in the four PCs components. In particular, PC1 explained 71.9% of the variance, with the main contribution in the Vis region, where Fe(II) and Fe(III) in primary minerals gave important insight into the chemical properties of such minerals. Charge transfer is photon stimulated by the high energy of the natural light, producing electron movement between nuclear centers of Fe with specific spectral information centered at 380 - 490 nm (PC1) (Sellitto et al., 2009). A Spearman test

showed a strong significant correlation of PC1 with soil variables related to andic characters (Ald, Fed, Feo, Alp, Fep, BS, PR and clay content) (Table 3, Spearman test >±0.5, p-level<0.05) The pronounced negative peak centered at 380 nm was most likely due to the strong absorption of poorly crystalline iron oxides.

PC2 explained 21.2 % of the variance. The most important loading weights were in the Vis range again, but with positive loadings, so that PC2 was directly proportional to the absorbance in the Vis range. A Spearman test showed a good significant correlation of PC2 with SOM and CEC (Table 2.3, Spearman test ≥±0.5, p-level<0.05). The loading weights of the PC2 show a broad positive peak centered at 400 nm, indicating a moderate influence of yellow-red visible wavelengths.

PC3 and PC4 account for a smaller amount of variation (respectively 5.5% and 0.7 % of explained variation). They probably explain the variability due to the large differences in particle size (PC4 was positively correlated with clay), parent material and other horizon peculiarities. PC4 was mostly correlated with the absorbance in the NIR region. There is also an explicit trend in PC3 and PC4 loading weights, focused at 670 nm (PC3) and increasing in contribution through to 1000 nm (PC4).

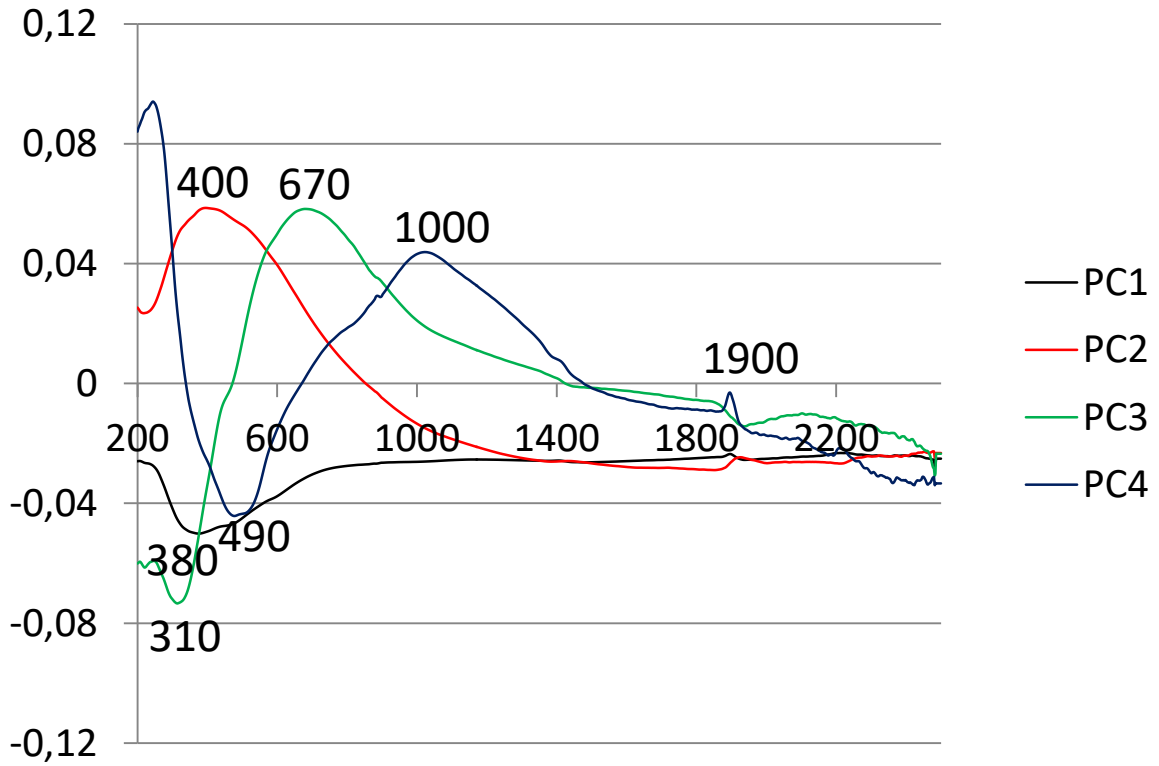


Figure 2.4: Loadings weights of the wavelengths in the first four PCs components

The discriminant analysis was performed on the following groups.

- Group A: Samples from profiles showing low andicity: EUR01 (Humi-Tephric Regosol (Eutric), EUR02 (Eutri-Humic Cambisol) of Italy and EUR18 (Molli-Vitric Andosol (Pachic) and EUR19 (Skeletic Umbrisol) of Hungary.
- Group B: Samples from the Fulvi-Silandic Andosol (Dystric) EUR03 and EUR04 of Italy
- Group C: Samples from the Vitric Andosols EUR07 (Orthidystri-Vitric Andosol), EUR08 (Dystri-Vitric Andosol), and EUR09 (Umbri-Vitric Andosol (Pachic and Orthidystric) of Iceland
- Group D: Samples from the Silandic Andosols EUR05 (Hyperdystri-Silandic Andosol) of Portugal; EUR10 (Umbri-Silandic Andosol (Hyperdystric) of Spain, EUR16 (Umbri-Silandic Andosol (Endoskeletic and Endoeutric) and EUR17 (Aluandi-Silandic Andosol (Umbric and Acrudoxic) of France
- Group E: Samples from the Silandic Andosols EUR06 (Hydri-Silandic Andosol (Umbric and Acroxic) of Portugal.

Group E included samples with the highest andic degree, namely with the highest values of Alo+1/2Feo (mean value 9.3%), pHNaF (mean value 11.1) and PR (mean value 99.8%). It was followed by the samples of the group D, that showed lower values especially for Alo+1/2Feo (mean Alo+1/2Feo value 5.3%; mean pHNaF value10.6 and PR 96.5%). The group A included samples with the lowest degree of andicity: the mean value of Alo+1/2Feo, was lower than 2% (mean value 1.5%), the mean values of PR and pHNaF were much lower than the threshold of andic properties (respectively mean PR value was 23% against 70% and pHNaF 8.0 against 9.5). The groups B and C included samples with a similar degree of andicity (mean Alo+1/2Feo value 3.1% and 2.5%; mean pHNaF value 9.5 and 9.1; mean PR value 80.3% and 79.7% respectively), but they mostly differed for the Silandic properties (mean Sio 1.33% in B vs 0.85% in C). The figure 2.5 a, b, c, d showed the mean values of andic parameter for each group.

According to a Spearman correlation test, the first seven PCs correlated with andic parameters were PC1, PC 3, PC 5, PC 7, PC 8, PC 10, and PC 14. The discriminant analysis was performed on this set of seven PCs and resulted to well classify samples according to the five groups: more than 86% of the samples were correctly classified, supporting the high sensitivity of VisNIR spectra to the level of expression of andic properties (figure 2.6). Actually, NIR spectra had already proved to offer an integrated vision of soil conditions (Cohen et al., 2004; Zornoza et al., 2008) and in this case a synthesis of andic parameters.

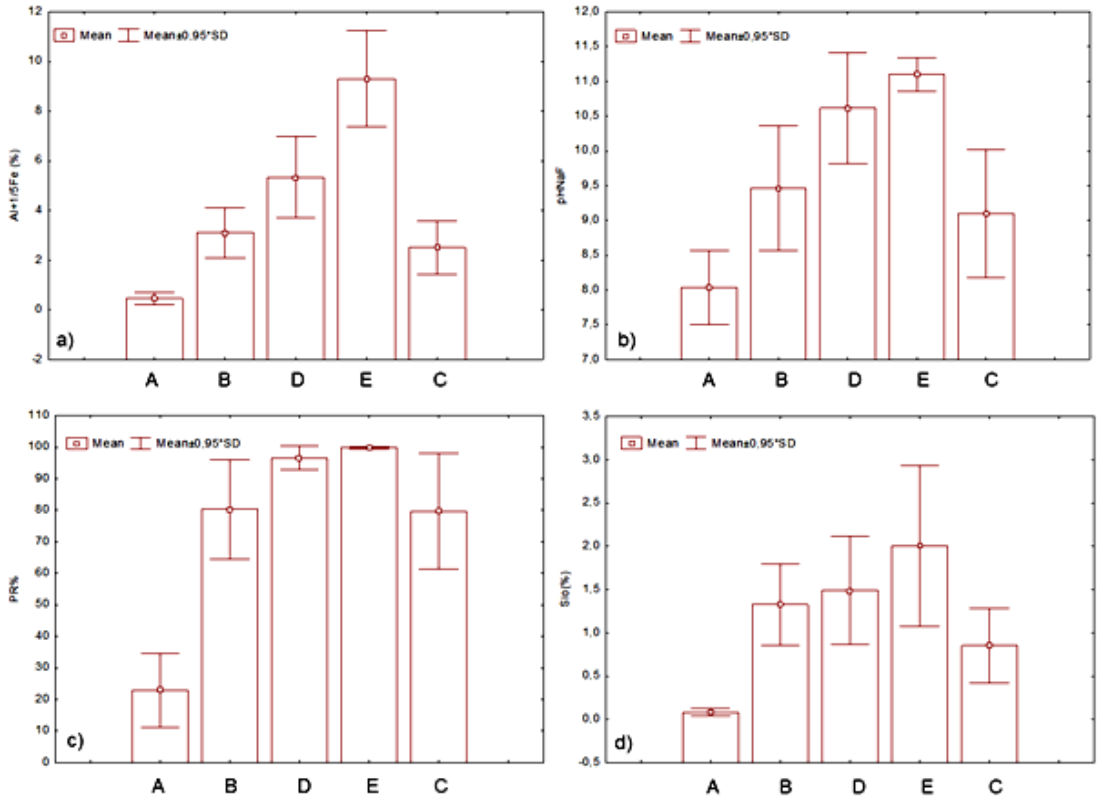


Figure 2.5: Andic parameters of the andic groups. a) mean values and standard deviation of $Al_o+1/2Fe_o$; b) mean values of pHNaF; c) mean value and standard deviation of PR; d) mean value and standard deviation of Sio.

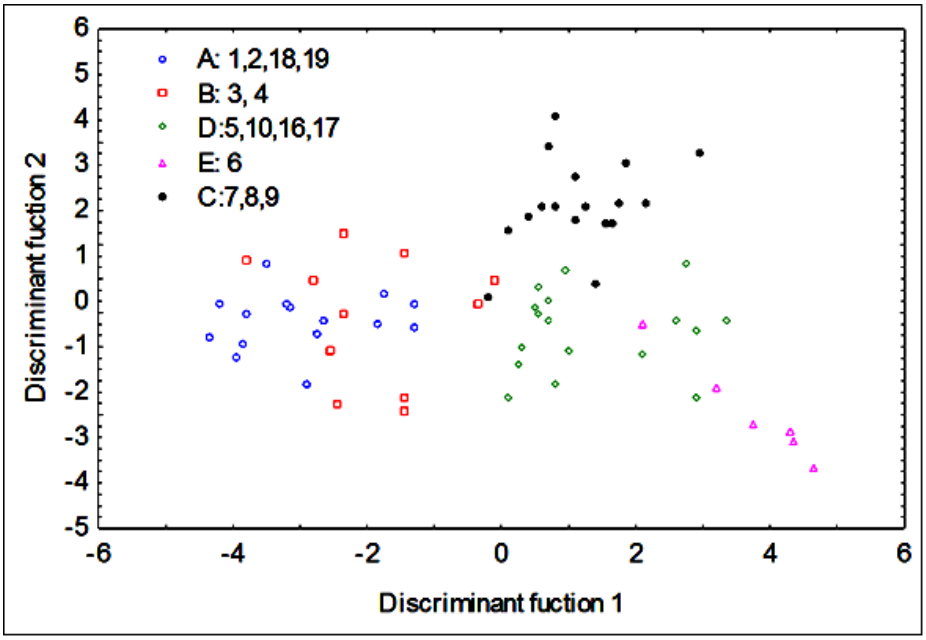


Figure 2.6. Discriminant analysis of the soil samples grouped by andic degree. The biplot axes represent the first two dimensions that provide maximum separation among the groups.

Table 2.3: Spearman Rank Order Correlations between Principal components (PC1…4) and soil properties. Marked correlations (*) are significant at p <0.05.

Parameter	PC1	PC2	PC3	PC4
Al_p	-0,63*	-0,17	-0,22	0,10
Fe_p	-0,62*	-0,20	-0,24	-0,03
Al_d	-0,59*	-0,15	-0,44*	0,15
Fe_o	-0,58*	-0,07	-0,48*	0,18
PR	-0,57*	-0,01	-0,41*	0,15
Clay	-0,54*	-0,15	0,20	0,32*
Fe_d	-0,52*	-0,15	-0,45*	0,23
$Al_o+1/2Fe_o$	-0,50*	0,02	-0,36*	0,25*
Al_o	-0,50*	0,05	-0,30*	0,27*
pH-NaF	-0,49*	-0,04	-0,28*	0,21
TOC	-0,46*	-0,53*	0,00	0,18
CEC	-0,46*	-0,49*	-0,07	0,18
Allophane	-0,41*	0,09	-0,31*	0,27*
Al_o/Si_o	-0,35*	-0,15	0,24	0,07
Si_o	-0,29*	0,09	-0,39*	0,30*
$(Al_o-Al_p)/Si_o$	-0,21	-0,02	0,24	0,16
pH-KCl	-0,04	-0,11	-0,19	0,31*
BS	0,60*	-0,16	0,46*	0,00

2.3.4 Prediction of andic soil properties using PLSR and SVM

Andic and main soil properties were estimated from the soil absorbance spectra in the VisNIR by PLSR. The best PLSR models were obtained with 5 or 6 factors (lowest RMSE). Five factors were used for the variables Fed, Ald, Sio, Alo, as a larger number of factors did not improve the prediction.

SMV models were carried out on the 7 or 6 PCs that showed a Spearman correlation with the target variables with a p-value <0.15. A lower number of PCs were used only for Ald (5 PCs), Fed (4 PCs), and Alp (4 PCs), since no other PCs were correlated. PC1 was adopted as

predictive variable in all the models, followed by PC3 (Tab. 2.3). PC2 instead was never adopted as predictor, since its strong correlation with the OC content (Tab. 2.3).

Some scatterplots of predicted vs. observed soil properties with both PLRS and SWM are presented in Figure 2.7 and 2.8. The selected scatterplots were showed for: Ald and Fed, which were the best predictable variables, Alo+1/2Feo, for its diagnostic importance in identifying the order of Andisol and as an example of different predictive power between the compared algorithms; PR because of its particular distribution of samples.

The results about the prediction accuracy and model parsimony from PLSR and SVM are reported in Table 2.4. The best results were obtained for Ald and Fed, which showed a good predictability both with SVM and PLSR (1.8<RPD<1.9; R^2 0.7), PLSR slightly outperforming SVM. Good results were also obtained also for Alo, Feo, Sio, Alo+1/2Feo, Alp and Alo-Alp, but only with SVM (1.5<RPD<1.8; 0.6<R^2<0.7), while PLRS was not able to predict them with the same accuracy (1.2<RPD<1.5, 0.3<R^2<0.5). The other andic parameters were not predicted within a threshold of acceptability, neither with PLSR nor with SVM. TOC was also poorly predicted (PLSR: R^2 = 0.40 and RPD = 1.26; SVM: R^2 = 0.40 and RPD = 1.24), while clay allowed only for a qualitative prediction (PLSR: R^2 = 0.50 and RPD = 1.4; SVM: R^2 = 0.55 and RPD = 1.49).

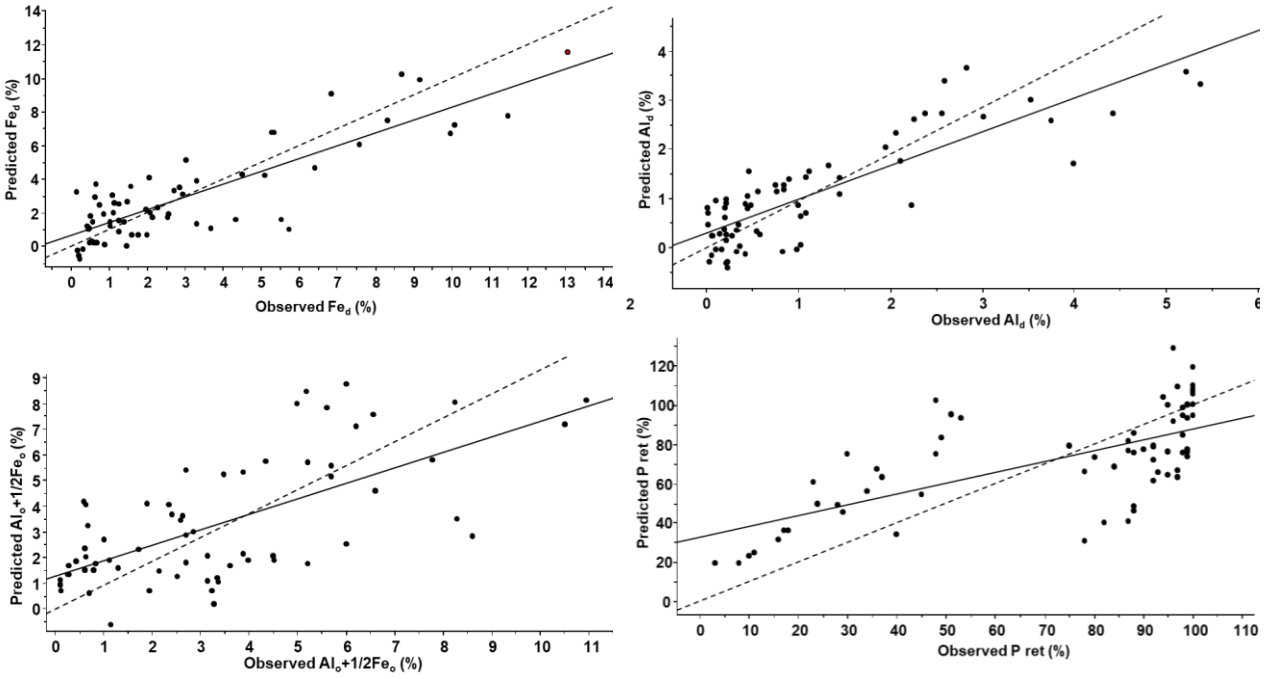


Figure 2.7. Scatterplots of observed (x-axis) vs. predicted values (y-axis) for validation dataset of some soil parameters (starting from the top-left: Fe_d; Al_d; Al_o+1/2Fe_o; P ret) ; the predictions were performed by Partial Least Squares Regression (PLSR). Full line is regression line, dashed line is target line.

Table 2.4. Comparison of predictions using PLSR and SWM algorithms.

		Al_d	Fe_d	Alloph.	Al_o	Fe_o	Si_o	Al_p	Fe_p	PR	Al_o+ 1/2Fe_o	Al_o- Al_p	TOC	Clay
Observed mean		*1.137*	*2.94*	*9.39*	*2.807*	*1.266*	*1.01*	*0.615*	*0.505*	*72.34*	*3.441*	*2.097*	*6.84*	*17.4*
Observed SD		*1.301*	*3.08*	*10.56*	*2.411*	*1.183*	*0.804*	*0.722*	*0.88*	*31.64*	*2.824*	*1.887*	*6.898*	*8.02*
MEAN	SVM	1.003	2.837	8.49	2.671	1.177	0.938	0.551	0.367	71.37	3.248	2.031	6.618	17.088
	PLSR	1.056	2.867	9.54	2.788	1.212	1.001	0.601	0.449	72.39	3.327	2.213	7.458	17.395
SD	SVM	0.911	2.292	6.03	2.069	0.916	0.598	0.451	0.48	25.52	2.301	1.607	5.393	6.102
	PLSR	1.068	2.767	7.90	2.005	0.804	0.521	0.585	0.47	26.02	2.458	1.691	7.225	8.103
RMSEP	SVM	0.763	1.665	8.026	1.385	0.69	0.532	0.482	0.654	23.5	1.589	1.271	5.423	5.377
	PLSR	0.702	1.632	9.287	1.826	0.933	0.694	0.548	0.85	23.88	2.088	1.596	5.716	5.696
R^2	SVM	0.676	0.715	0.428	0.67	0.663	0.563	0.573	0.487	0.458	0.684	0.552	0.396	0.545
	PLSR	0.708	0.718	0.26	0.445	0.375	0.261	0.438	0.103	0.446	0.479	0.4	0.404	0.501
Bias	SVM	0.882	0.965	0.904	0.951	0.93	0.929	0.896	0.728	0.987	0.944	0.969	0.967	0.979
	PLSR	0.929	0.975	0.995	0.993	0.957	0.99	0.977	0.889	1.001	0.967	1.009	1.026	0.995
SEP	SVM	1.068	1.886	8.035	1.613	1.092	1.013	0.966	0.877	23.66	1.76	1.57	5.511	5.444
	PLSR	1.106	1.878	9.406	2.086	1.308	1.211	1.117	1.198	24.08	2.27	1.913	5.886	5.81
RDP	SVM	1.705	1.85	1.316	1.741	1.716	1.509	1.497	1.345	1.346	1.778	1.484	1.272	1.491
	PLSR	1.852	1.887	1.127	1.32	1.268	1.157	1.317	1.036	1.325	1.353	1.254	1.264	1.422

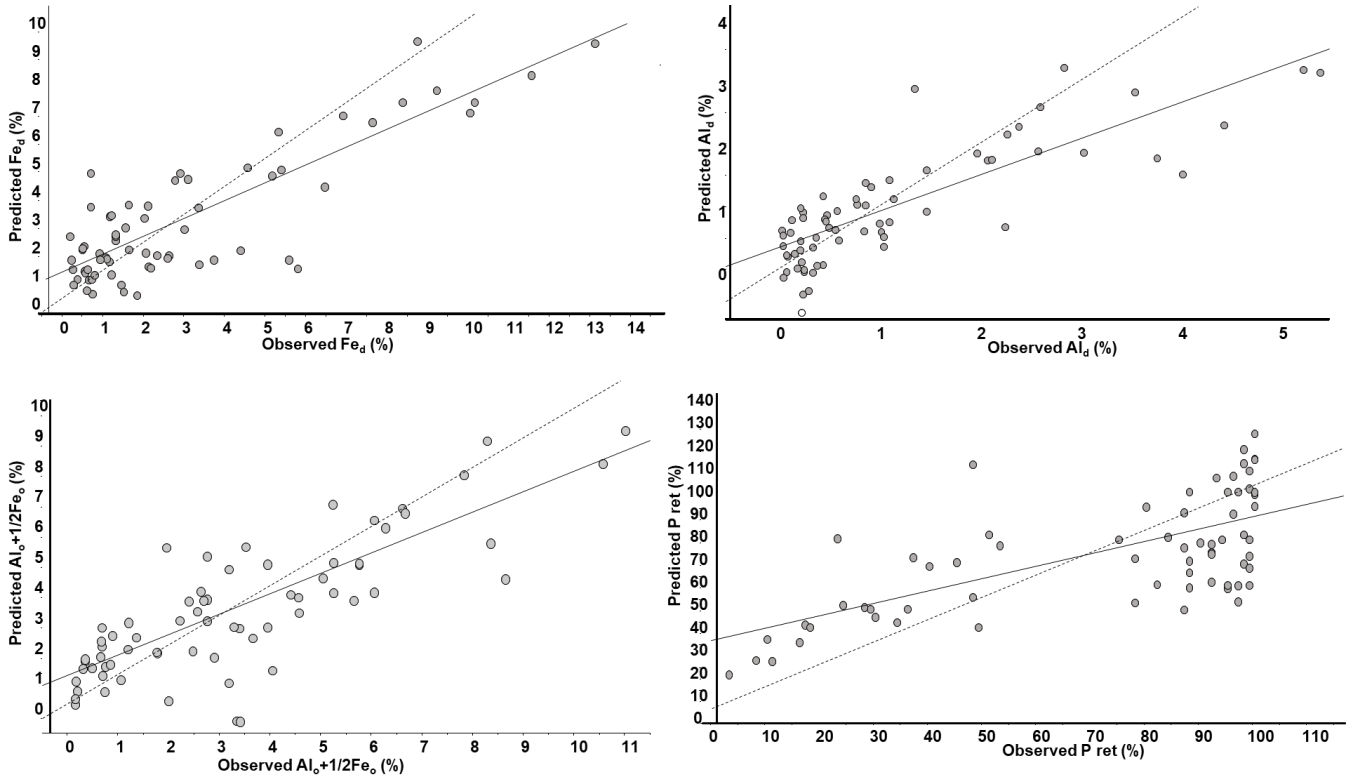


Figure 2.8: Scatterplots of observed (x-axis) vs. predicted values (y-axis) for validation dataset of some soil parameters (starting from the top-left: Fe_d; Al_d; Al_o+1/2Fe_o; P ret); the predictions were performed by Supported Vector Machine (SVM). Full line is regression line, dashed line is target line.

2.4. Discussion

2.4.1 Estimation of andic properties with VisNIR

According to our results, the use of VisNIR could be considered statistically suitable for the estimation of free Al and Fe forms, expressed as Ald and Fed, even if an error occurs, so that the results may be useful at least for a rough qualification of the presence of these Al and Fe forms. The same might be affirmed for the forms of Al bounded in organic complex (Alp) and for Al in allophane and imogolite expressed as Alo-Alp, but findings resulted at a minimum acceptance threshold.

As regards the elements extracted with the ammonium-acid oxalate, they represent the compounds with a low crystalline order of iron, aluminum and silicon and not a unique form. In particular, Alo corresponds with (i) Al bonded in organic complexes, (ii) non-crystalline hydrous oxides and (iii) allophane, imogolite and others. Therefore, the quantity extracted as Alo cannot be ascribed to a specific fraction, as the total non-crystalline (Alo) corresponds to a group of compounds which can vary between soils. Its prediction through models may therefore be particularly difficult. For this reason and considering the poor predictive ability of PLRS, even if a good RPD was obtained with SVM (1.5<RPD<1.8), the predicted Alo, Feo, Alo+1/2Feo, and

Sio by DRS in VisNIR could be considered as a preliminary result. Further studies on larger databases are needed to confirm our findings. On the other hand, the prediction of the forms of Fe bounded in organic complex (Fep), could not be estimated from VisNIR, because of the very poor prediction both with SVM and PLSR. As a whole, the partially positive results on the total non-crystalline forms and on the forms bounded in organic complex could encourage the expectation for a predictability that might be improved in the future with a wider database and more specific study.

As regards SOM, many previous studies have pointed to TOC as a very well high predictability variable by spectroscopy, also on Andosols. Some difficulties in prediction TOC in Andosols may occur due to the existence of overlapping absorption features, so that absorptions related to one soil component can be masked, distorted, or shifted where other soil components vary. For instance, Adar et al., (2014) found that spectral variations related to changes in iron oxide content may cancel variations in absorptions due to organic matter. Bellino et al. (2015) found that SOM can be properly predicted using few observations of andic soil profiles and a combination of calibration techniques (Supervised Principal Component regression/ Least Absolute Shrinkage and Selection Operator) on UV-Vis-NIR spectra in the range 200-2500 nm. Zhao et al. (2012) investigated the prediction of SOM content in Andosols using Portable Hyper-spectral Camera and they showed that combined 360-1010 nm and 900-1700 nm spectrographs produced acceptable results. In a recent study carried out in a coffee agroforestry system on Andosols, Kinoshita et al. (2016) highlighted that assessing topsoil organic carbon by Vis- NIR spectroscopy gave better results than other smaller datasets of auxiliary variables from laboratory analyses. Bonett et al. (2016) found a shared model to predict SOM both in Andosols and other soil types (Vertisols and Oxisols), as in these soil types the properties were manifested in similar spectral regions, but with different levels of reflectance. However, the results of this work indicate that TOC estimation showed one of the worst results. It was probably due to the higher TOC variability caused by the presence in both C, B and A horizons, differently from the database of other mentioned studies (Kinoshita et al., 2016 studied topsoils; Zhao et al., 2012, studied the first 0-2.5cm).

2.4.2 Model comparison

The findings about the comparison of predictions using PLSR and SVM varied greatly depending on the predicted andic soil parameter and on the model. In the cases of free Al and Fe forms a good prediction was obtained from both, with a slight better result for PLSR. PR and the Al and Fe forms bounded in organic complex, the allophane, were poorly predicted by the use of both the type of models, even if results from SVM reached a RPD of 0.5 for Alp. Conversely, our findings suggest how SVM can be considered a tool stronger than PLSR to predict as a whole the poorly crystalline forms of Al, Fe and Si, and the relative Alo +1/2 Feo index. Therefore, our results partially disagreed with what Viscarra Rossel and Behrens (2010) found about other soil properties (SOC, clay content and pH), for which PLSR produced competitive predictions to SVM.

All the models were affected from error in the precision (SEP) and in the BIAS, but in general the error was mostly due to the lack of precision, both for PLRS and for SVM. Best predicted parameters (Ald and Fed) were estimated with similar SEP and BIAS of models performed with SMV and PLSR. On the other side, when SVM outperformed PLSR, the SEP most differentiated PLSR from SVM.

2.5. Conclusion

The findings of this study show that DRS can be used for estimating some andic parameters within a good or acceptable predictability according to RPD, RMSE and R^2. In most of cases, SVM was able to pick up the relationship between spectra and andic properties better than PLSR, since the lack of precision (SEP) affected PLSR more than SVM. Both SVM and PLSR estimate free Al and Fe oxydroxides evaluated with the extracted dithionite-citrate extractable (Ald and Fed) in a good way. However, good estimations were obtained with SVM also for Al and Fe total non-crystalline (Alo and Feo), the index acid oxalate extractable for Alo+1/2Feo, but in a lower measure for Si total non-crystalline (Sio), oxalate Al extracted from Al-humus complex (Alp), and Al included in allophane and imogolite expressed as Alo-Alp.

In conclusion, our results highlight that VisNIR information could be used as a preliminary easy method to distinguish Andosols from other soils and to highlight different degrees of andic properties. According to the predictive calibration models, it might be more difficult to use VisNIR to pick up differences between silandic and aluandic properties, since no one estimations exceeded the threshold to consider the predictions excellent (RPD always <2.0). Nevertheless, based on the encouraging results achieved (1.5>RPD>1.9) further studies, with a larger database, could lead to more reliable models and return quantitatively accurate estimations for several investigated andic parameters.

The main outcome of this research is that VisNIR spectra contain enough information to recognize the volcanic origin of soil profiles and their level of expression of andic traits. Using VisNIR in the laboratory, good predictions can be achieved for many diagnostic criteria of andic horizons, which can also be used to differentiate soil horizons. Using VisNIR spectroscopy could be also used by remote sensing. to predict andic properties and classify volcanic soils of wide territories through a single flight campaign.

Acknowledgements

Firstly I thank the co-authors of this paper, Erika Di Iorio, Edoardo A.C. Costantini and Claudio Colombo. I also thank: the Working Group for Soil Description and Sampling of COST 622 (François Bartoli, Angelo Basile, Martine Gérard, Toine Jongmans, Folkert van Oort and Fabio Terribile); P. Buurman for coordinating the COST 622 chemical data base and organic carbon and soil chemical extraction data were kindly provided by E. Garcia-Rodeja of the University of Santiago, Spain.

References

Arnalds, O., Bartoli, F., Buurman, P., Oskarsson, H., Stoops, G., Garcia-Rodeja, E., 2007. ED. 2 Soils of Volcanic Regions in Europe. Springher Verlag. pp. 644

Adar, S.Y., Shkolnisky, Y., Ben Dor, E., 2014. Change detection of soils under small-scale laboratory conditions using imaging spectroscopy sensors. Geoderma 216, 19e29.

Bédidi, A., Cervelle, B., Madeira, J., Pouget, M., 1992. Moisture effects on visible spectral characteristics of lateritic soils. Soil Sci. 153,129–141.

Bellino, A., Colombo, C., Iovieno, P., Alfani, A., Palumbo, G., Baldantoni, D. 2015. Chemometric technique performances in predicting forest soil chemical and biological properties from UV-Vis-NIR reflectance spectra with small, high dimensional datasets. iForest-Biogeosciences and Forestry, 9(1), 101-108.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC Trends in Analytical Chemistry, 29(9), 1073-1081.

Ben-Dor, E., Banin, A., 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Science Society of America Journal 59, 364–372.

Ben-Dor, E., Heller, D., Chudnovsky, A., 2008. A novel method of classifying soil profiles in the field using optical means. Soil Science Society of America Journal 72, 1113–1123.

Bishop, J. L. Ethbrampe, E. B.; Bish, D. L., Abidin, Z. L., Baker, L. L., Matsue, N., Henmi, T., 2013. Spectral and Hydration Properties of Allophane and Imogolite. Clays and Clay Minerals 61(1), 57-74.

Bonett, J.P., Camacho-Tamayo, J.H., Vélez Sánchez, J.E., 2016. Estimating soil properties with mid-infrared spectroscopy. Rev. U.D.C.A Act. & Div. Cient. 19(1), 55-66.

Brown, D.J., Bricklemyer, R.S., Miller, P.R. 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma 129(3–4), 251–267.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M.D., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132, 273–290.

Buurman, P., van Lagen, B., and Velthorst, E.J., 1996. Manual for Soil and Water Analysis. Backhuys Publishers, Leiden, The Netherlands. pp 314.

Buurman, P., García-Rodeja, E., Martínez Cortizas, A., van Doesburg, J.D.J. 2003. Stratification of parent material in European volcanic and related soils studied by laser diffraction grain-sizing and chemical analysis. Catena, 56, 127-144.

Calvin, W.M., Littlefield, E.F., Kratt, C., 2015. Remote sensing of geothermal-related minerals for resource exploration in Nevada. Geothermics 53, 517–526.

Cañasveras, J.C., Barrón, V., del Campillo, M.C., Torrent, J., Gómez, J.A., 2010. Estimation of aggregate stability indices in Mediterranean soils by diffuse reflectance spectroscopy. Geoderma 158, 78-84.

Cécillon, L., Cassagne, N., Czarnes, S., Gros, R., Vennetier, M., Brun, J.J., 2009. Predicting soil quality indices with near infrared analysis in a wildfire chronosequence. Science of the Total Environment 407, 1200–1205.

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.

Chong, I.G., Jun, C. H., 2005. Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems, 78, 102–112.

Clark, R.N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. In: Rencz, N. (Ed.), Remote Sensing for the Earth Sciences: Manual of Remote Sensing. John Wiley & Sons, New York, pp. 3–52.

Cohen, M.J., Prenger, J.P., DeBusk, W.F., 2005. Visible-near infrared reflectance spectroscopy for rapid, non-destructive assessment of wetland soil quality. J. Environ. Qual. 34, 1422–1434.

Colombo, C., Sellitto, V.M., Palumbo, G., Di Iorio, E., Terribile, F., Schulze D.G. 2014. Clay formation and pedogenetic processes in tephra-derived soils and buried soils from Central-Southern Apennines (Italy). Geoderma 213,346-356.

Cozzolino, D., Moron, A., 2003. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. Journal of Agricultural Sciences 140, 65–71.

Dahlgren, R.A., Saigusa, M., Ugolini F.C., 2004. The nature, properties and management of volcanic soils. Adv. Agron. 82, 113–182.

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.

Ge, Y., Morgan C.L.S., Grunwald, S., Brown D.J., Sarkhot D.V., 2011. Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. Geoderma 161, 202–211

García-Rodeja, E., Nóvoa, J.C., Pontevedra, X., Martínez, and A., Buurman, P., 2004. Aluminium fractionation of European volcanic soils by selective dissolution techniques. Catena 56:155-183.

García-Rodeja, E., Nóvoa, J.C., Pontevedra, X., Martínez-Cortizas, A., Buurman, P., 2007. Aluminium and iron fractionation of European volcanic soils by selective dissolution techniques. In: Arnalds, Ó., Bartoli, F., Buurman, P., Óskarsson, H., Stoops, G., García- Rodeja, E. (Eds.), Soils of Volcanic Regions in Europe. Springer, Berlin, pp. 325–351.

Gomez, C., Le Bissonnais, Y., Annabi, M., Bahri, H., Raclot, D., 2013. Laboratory Vis–NIR spectroscopy as an alternative method for estimating the soil aggregate stability indexes of Mediterranean soils. Geoderma 209, 86-97.

Grunwald, S., Yu, C., Xiong, X., 2014. Transferability and scaling of soil total carbon prediction models in Florida. PeerJ PrePrints 2:e494v1 https://doi.org/10.7287/peerj.preprints.494v1

IUSS, 2014. Working Group World Reference Base for Soil Resources (WRB)

Karatzoglou, A., Smola, A., Hornik, K., 2008. Kernlab: Kernel-based Machine Learning Lab.(At: http://cran.r-project.org/web/packages/kernlab/index.html. Accessed: 08/04/2017)

Kinoshita, R., Roupsard, O., Chevallier, T., Albrecht, A., Taugourdeau, S., Ahmed, Z., Van Es, H. M., 2016. Large topsoil organic carbon variability is controlled by Andisol properties and effectively assessed by VNIR spectroscopy in a coffee agroforestry system of Costa Rica. Geoderma 262, 254-265.

Kodaira, M., Shibusawa, S., 2013. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. Geoderma 199, 64-79.

Jain, P., Kulis, B., Davis, J. V., Dhillon, I. S., 2012. Metric and kernel learning using a linear transformation. Journal of Machine Learning Research, 13, 519-547.

Iamarino, M., Terribile, F., 2008. The importance of andic soils in mountain ecosystems: a pedological investigation in Italy. Eur. J. Soil Sci. 59, 1284–1292.

IUSS Working Group., 2014. World reference base for soil resources 2014 international soil classification system for naming soils and creating legends for soil maps. FAO, Rome.

Madeira, M., Auxtero, E., and Sousa, E., 2003. Cation and anion exchange properties of andisols from the Azores, Portugal, as determined by the compulsive exchange and the ammonium acetate methods. Geoderma 117: 225–241.

Matus, F., C. Rumpel, R., Neculman, M., Panichini, M.L., Mora, L. 2014. Soil carbon storage and stabilisation in andic soils: A review. Catena 120, 102–110.

McDowell, M.L., Bruland, G.L., Deenik, J.L., Grunwald, S., Knox, N.M., 2012a. Soil total carbon analysis in Hawaiian soils with visible, near-infrared and mid-infrared diffuse reflectance spectroscopy. Geoderma 189–190, 312–320.

McDowell, M.L., Bruland, G.L., Deenik, J.L., Grunwald, S., 2012b. Effects of Subsetting by Carbon Content, Soil Order, and Spectral Classification on Prediction of Soil Total Carbon with Diffuse Reflectance Spectroscopy. Applied and Environmental Soil Science 2012, 14.

Minasny, B., McBratney, A.B., 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. Chemom. Intell. Lab. Syst. 94 (1), 72–79.

Mouazen, A.M., Kuang, B., de Baerdemaeker, J., Ramon, H., 2010. Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. Geoderma 158, 23–31.

Nanni, M.R., Demattê, J.A.M., 2006. Spectral reflectance methodology in comparison to traditional soil analysis. Soil Science Society of America Journal 70, 393–407.

Nanzyo, M., Dahlgren, R., and Shoji, S., 1993. Chemical characteristics of volcanic ash soils. In Shoji, S., Nanzyo, M., and Dahlgren, R.A., eds., Volcanic Ash Soils. Genesis, Properties and Utilization. Developments in Soil Science 21. Amsterdam: Elsevier, pp. 145–187.

Nocita, M., Stevens A., Van Wesemael B., Aitkenhead M., Bachmann M., Barthès Bernard, Ben Dor E., Brown D.J., Clairotte M., Csorba A., Dardenne P., Demattê J.A.M., Genot V., Guerrera C., Knadel M., Montanarella L., Noon C., Ramirez-Lopez L., Robertson J., Sakai

H., Soriano-Disla J.M., Shepherd K.D., Stenberg B., Towett E.K., Vargas R., Wetterlind J., 2015. Soil spectroscopy: an alternative to wet chemistry for soil monitoring. Advances in Agronomy 132, 139-15.

Parfitt, R.L., Kimble, J.M., 1989. Conditions for formation of Allophane in soils. Soil Science Society of America Journal 53, 971–977.

Priori, S., Fantappie, M., Bianconi, N., Ferrigno, G., Pellegrini, S., Costantini, E.A.C., 2016. Field-Scale Mapping of Soil Carbon Stock with Limited Sampling by Coupling Gamma-Ray and Vis-NIR Spectroscopy. Soil Science Society of America Journal 80, (4) 954-964.

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

Sellitto, V.M, Barrón, V., Palumbo, G., Colombo C., 2007. Application of Diffuse Reflectance Spectroscopy (DRS) to study European Volcanic Soils: a preliminary examination. In Soils of Volcanic Regions in Europe. (Ed). Springer, Berlin, Heidelberg, New York. pp. 437-452.

Sellitto, V.M, Fernandes, R.B.A, Barron, V, Colombo C., 2009. Comparing two different spectroscopic techniques for the characterization of soil iron oxides: Diffuse versus bi-directional reflectance. Geoderma 149, 2-9.

Sherman, D.M., Waite, T.D., 1985. Electronic spectra of Fe^{3+} oxides and oxyhydroxides in the near infrared to ultraviolet. American Mineralogist 70, 1262–1269.

Shepherd, K.D., Walsh., M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66:988–998.

Shoji, S., Toyoaki, I., Saigusa, M., Yamada, I., 1985. Properties of nonallophanic Andosols from Japan. Soil Science 140, 264 – 277.

Shoji, S., Nanzyo, M., Dahlgren, R.A., 1993. Volcanic Ash Soils. Genesis, Properties and Utilization. Developments in Soil Science, vol. 21. Elsevier, Amsterdam. pp 288.

Stevens, A., Nocita, M., Toth G., Montanarella, L., van Wesemael, B., 2013. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. PLoS ONE 8(6): e66409. doi:10.1371/journal.pone.0066409

Stone, M., Brooks, R., 1990. Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. Journal of the Royal Statistical Society, Series B 52(2), 237-269.

Takahashi, T., Shoji, S., 2002. Distribution and Classification of Volcanic Ash Soils. Global Environmental Research 6, 83-97.

Torrent, J., Barrón, V., 1993. Laboratory measurements of soil color: theory and practice. In J.M. Bigham and E.J. Ciolkosz (ed.) Soil color. SSSA Spec. Publ. 31. SSSA, Madison, WI, pp. 21–33.

Ugolini, F.C., Dahlgren, R.A., 2002. Soil Development in Volcanic Ash. AIRIES, Volcanic Ashes and their Soils. Association of International Research Initiatives for Environmental Studies, Tokyo pp. 69–81.

Vapnik, V., 2000. The nature of statistical learning theory. Second. Springer, New York, USA

Viscarra Rossel, R., 2008. ParLeS: Software for chemometric analysis of spectroscopic data. Chemometrics and Intelligent Laboratory Systems 90, 72–83.

Viscarra Rossel, R.A., Cattle, S., Ortega, A., Fouad, Y., 2009. In situ measurements of soil colour, mineral composition and clay content by Vis-NIR spectroscopy. Geoderma 150, 253–266.

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

Wada, K., S. Aomine., 1973. Soil development on volcanic material during the Quaternary. Soil Sci. 116:170-177.

Wada, K., 1985. The distinctive properties of Andosols. Advances of Soil Science, vol. 2. Springer-Verlag, New York, 2, 173–228.

Wang, Y.B., Huang, T.Y., Liu, J., Lin, Z.D., Li, S.H., Wang, R.J., Ge, Y.J., 2015. Soil pH value, organic matter and macronutrients contents prediction using optical diffuse reflectance spectroscopy. Comput. Electron. Agric. 111, 69–77.

Williams, P., 1987. Variables affecting near-infrared reflectance spectroscopic analysis. In Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists, St. Paul, Minnesota. pp. 143–167.

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

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems 58(2), 109-130.

WRB-FAO (World Reference Base for Soil Resources). 2014. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports, 106). Rome: Food and Agriculture Organization of the United http://www.fao.org/3/a-i3794e.pdf Nations.

Zornoza, R., Guerrero, C., Mataix-Solera, J., 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 and Biochemistry, 40(7), 1923-1930.

Zhao, T., Sakai, K., Higashi, T., Komatsuzaki, M., 2012. Assessing Soil Organic Carbon Using Portable Hyper-spectral Camera in Andisols. J. Agric Sci. Appl. 1(4), 131-137.

CAP.3

MONITORING SOIL ORGANIC CARBON BY NEAR INFRARED SPECTROSCOPY (NIR)

3.1. Introduction

The organic carbon (OC) content in the soil is considered one of the most important properties in assessing soil quality (Andrews *et al.*, 2004) for its several positive effects, namely, on soil structure, cation exchange capability and water-holding capacity. Moreover, the soils ability in the carbon sequestration and mitigation of greenhouse gases in the atmosphere makes growing the interest on soil carbon content (Viscarra Rossel *et al.*, 2011). As global soils contain 2 to 3 times more carbon than the atmosphere (URL:http://4p1000.org/understand), a relatively small increase in the stock can significantly improve the mitigation of greenhouse gases emissions. The idea was launched in December 2015, during COP21 by the French Minister of Agriculture (Minasny et al., 2017). Now, the international research program '4 per mille Soils for Food Security and Climate' aims to compensate the global emissions of greenhouse gases by anthropogenic sources, by increasing the global soil organic matter stocks. An absolute increase of 0.4% or 4‰ per year shall be considered as satisfactory, since it represents the ratio between the global anthropogenic C emissions and the total soil OC stock (Minasny et al; 2017). This growth rate is intended to show that even a small increase in the soil carbon stock is crucial to improve soil fertility and agricultural production and to contribute to achieve an effect on climate change (URL: http://4p1000.org/understand).

An OC increase in the soil can be pursued by appropriate land management changes in cropping, grazing, horticultural and mixed farming systems, since the introduction of conservation practices can lead to sequester OC over conventional practices (West and Post, 2002; Lal, 2004). As a consequence, measuring, monitoring, and comparing the effect of the introduction of different innovative practices have acquired a strategic role to assess their effects on C stocks.

Conventional methods to monitoring the soil OC content involve field soil sampling, followed by sample preparation and laboratory wet analysis. These methods are time consuming and expensive. If a few samples are required for monitoring, the conventional methods may be not a problem. It could happen when a big difference occurs between the data that we are going to compare. However, the introduction of new agricultural practices leads to a very slow change in soil OC (Smith, 2004). Consequently, the OC differences that we need to measure can be very little, many samples are required, and conventional methods turn into a limitation for costs and time. Therefore, we need the development of alternative accurate and cost-efficient methods for measuring and monitoring the changes.

Different approach may be adopted to reduce costs for monitoring change in soil OC content.

One solution may consist in taking into account the cumulative effect of the OC increase in soil along the time (Minasny et al., 2017). The longer the time after the introduction of a new land use management, the bigger the expected OC increase. So, a long period of observation can reduce the number of samples necessary for monitoring. Some authors suggest to wait more than 5 to 10 years (Smith, 2004). But this represents only a partial solution to the problem. It may be useful to obtain result in a shorter time. Moreover, it may be that even when a big increase occurs, the variance is so high that a little number of samples could be not enough to detect differences. In fact, according to the power analysis, the difference to detect has an inverse correlation with the size number necessary to assure an accurate and reliable estimation of the OC increase, while the variance within the dataset has a direct correlation.

Another solution may consist in monitoring through data achievable in an inexpensive way. The idea is to find relationships between one set of measurements easy and cheap to acquire, and conventional laboratory OC measurements. Once good relationships are found, the OC contents can be predicted rapidly and with high accuracy from the cheaper measure, so that a large amount of data may be used to detect OC changes. Proximal sensing can provide this kind of economic information. The most suitable technique for measuring soil OC concentrations was recognized to be visible and near-infrared reflectance spectroscopy (England and Rossel, 2018). Indeed, many studies have already demonstrated that very good predictions can been obtained estimating soil physical, chemical, and biological properties by the use of diffuse reflectance measurements in the near-infrared region (NIR; 780–2500 nm; Sheppard et al., 1985) (Viscarra Rossel et al., 2006; Veronique et al., 2010; Shi et al.,2014). In particular, very good predictions have been obtained in modelling OC, (Viscarra Rossel et al., 2016, Reeves et al., 2002; Sørensen and Dalsgaard, 2005, Xie, 2011, Stevens et al., 2008; Morgan et al., 2009). Some authors proposed NIR for monitoring soil OC changes (Cécillon et al., 2009, Stevens et al., 2008; Nocita et al., 2015).

However, also predictive models need some efforts because they need to be calibrated with a certain number of laboratory measurements, in order to find the relationships of spectroscopic soil properties with the actual OC content. The number of laboratory measurements employed for the calibration or for the adaptation (or recalibration) of existing models represents a limitation for the cost-effective benefit by the use of NIR spectroscopy.

Moreover, even in the case of accurate models, the predictions are not completely free from error. When the error affects the ability to reproduce exactly difference in OC content between the studied populations and the variance of the populations, the error may reduce the efficiency in detecting the OC increase.

The first objective of our study was to find what type of NIR calibration was the best one for monitoring OC in the soil. Calibrations were carried out with the use of local datasets and also with the help of big general dataset available for Spain. With a view to limit the scenarios where NIR-based approach can assure a cost-effective benefit, the study focused on a way to quantify the effort in terms of number size for balancing the error in checking the change. The scenarios were mainly intended in terms of what is the range of the OC increase we expected to

be monitored. Finally a cost benefit analysis was carried out in the frame of each investigated scenario.

3.2. Materials and methods

3.2.1. Site

The soil samples used in this study were collected in two agricultural fields located in Tarazona de la Mancha (Albacete, Spain) in December 2017. Both fields are separated by less than 2 kilometres, and are sharing similar characteristics such as soil type (Calcixerept over old river terraces), meteorological conditions (mean temperature 14°C; mean annual rainfall 500 mm), slope (<2%), parent material (calcareous and fluvial deposits), or rotation crops (wheat, maize, etc.) and irrigation system (pivots on circular fields). The unique difference is that one of the fields changed its management from conventional to no-tillage 20 years ago (NT20), while the other remains under the conventional tillage (CON). The size of the CON field is approximately 60 hectares, whereas the NT20 field is 45 hectares, being both circular shape (Fig. 3.1).

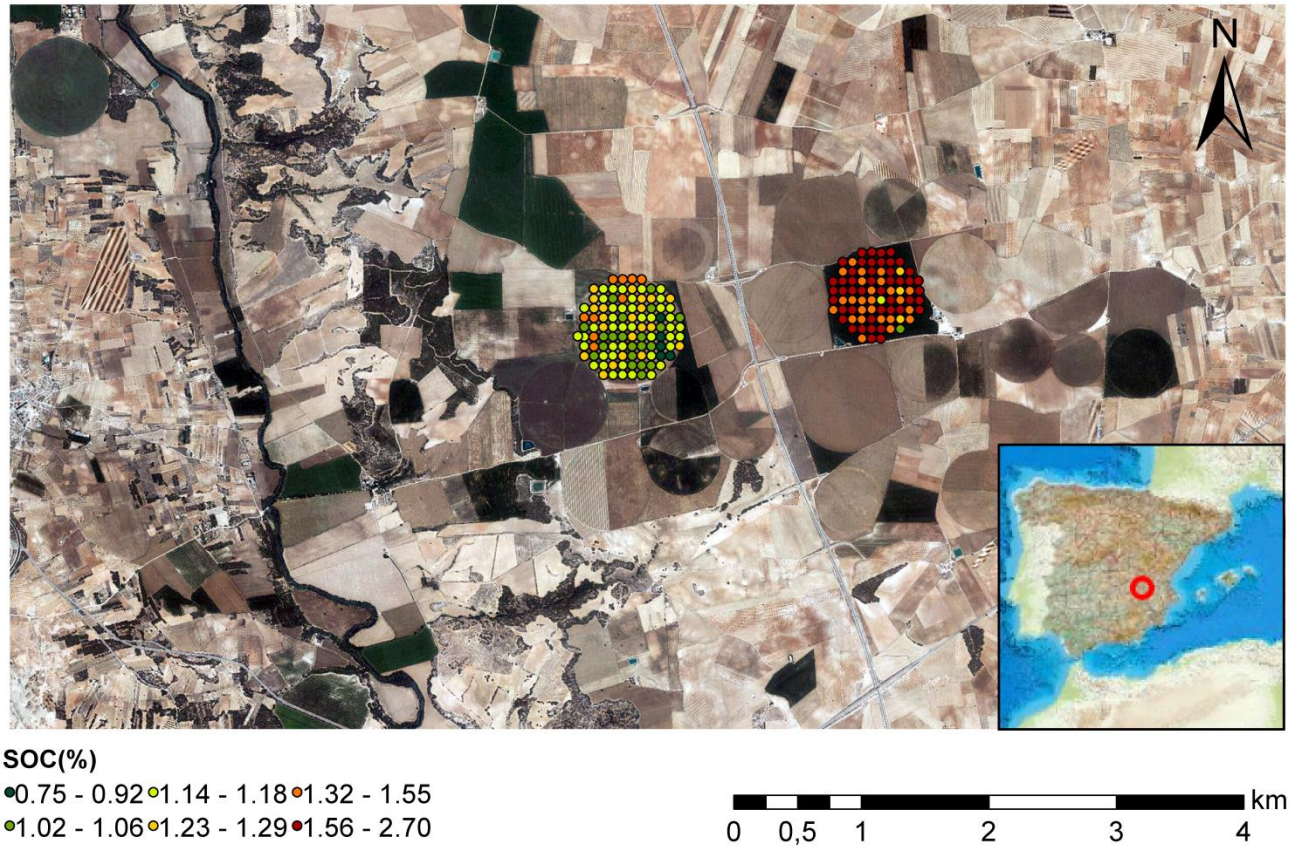


Figure 3.1: Map of the investigated area. The field on the left with lower SOC % values is the conventional one, the field on the right with higher SOC% values is under 20 years of no tillage (ortophoto image from http://www.ign.es/iberpix2/visor/).

3.2.2.- Samples and laboratory analysis

The samples were collected approximately in the intersections of a regular grid (75 meters) randomly located in each field. In that way, we collected 100 samples in CON and 76 samples in NT20. Each sample was composed by several (1-meter adjacent) subsamples collected with an auger from 0 to 10 cm depth. The soil samples were air-dried in laboratory (25°C) and sieved (<2mm). The SOC content in each soil sample was analysed with the Walkley-Black (1934) (WB) method in duplicate. The NIR spectra were obtained by Fourier Transform (FT)–NIR diffuse reflectance spectroscopy (MPA, Bruker Optik GmbH, Ettlingen, Germany). Each spectrum was composed by 64 scans, and two spectra per sample were acquired and then averaged. The x-scale of the spectra was transformed to nanometres (830–2630 nm) and resampled to 1-nm resolution.

Two samples from NT20 were discarded due to discrepancies between NIR and WB. For most of the experiments, we used the remaining 74 NT20 samples and a random selection of 74 CON samples.

2.3. Scenarios

We have established four temporal scenarios, defined by the time after the change in the land management from conventional to no-tillage: 5, 10, 15 and 20 years. The variation in SOC content (ΔSOC) after 20 years of the land management change was defined as ΔSOC20, and it was calculated using Equation 3.1:

$$\Delta SOC_{20} = \mu_{NT20} - \mu_{CON} \quad \text{[Eq. 3.1]}$$

where μ_{CON} is the mean SOC concentration of the 74 samples collected in CON and μ_{NT20} is the mean SOC concentration of the 74 samples collected in NT20. On the basis of these observed values, we have estimated the ΔSOC after 5 years (ΔSOC_5) 10 years (ΔSOC_{10}) and 15 years (ΔSOC_{15}). These values have been estimated using a simple linear interpolation, assuming i) that the ΔSOC follows a linear cumulative pattern, ii) the SOC contents in both fields was similar before the change, and iii) the SOC content in CON remains constant. Some authors have observed changes following non-linear patterns (e.g. sigmoid, asymptotic, etc….).

We assumed that the assumption of a linear change is enough for the purpose of this experiment. Similarly, the mean (pooled) variance (σ^2) at each temporal scenario has been interpolated from values measured in samples collected at CON (σ^2_{CON}) and NT20 (σ^2_{NT20}). These variances were identified as σ^2_{0-5y}, σ^2_{0-10y}, σ^2_{0-15y} and σ^2_{0-20y}, for the mean variance observed after 5, 10, 15 and 20 years of the change, respectively.

3.2.4. Activities carried out in the frame of each scenario

3.2.4.1. Samples size

The sample size (n) needed to detect the ΔSOC5, ΔSOC10, ΔSOC15 and ΔSOC20 by traditional approach was calculated using power analysis (Eq. 3.2), at α=0.05 and β=0.10:

$$n = \frac{(t_{(\alpha/2)} + t_{(1-\beta)})^2 \times 2\sigma^2}{\Delta SOC^2}$$ [Eq. 3.2]

where t(α/2) is 1.96 (at α=0.05), t(1-β) is 1.282 (at 1-β=0.90), and σ^2 is the pooled variance in each scenario. In this way, the sample size needed to detect the ΔSOC5, ΔSOC10 and ΔSOC15, ΔSOC20 was identified as n_{5y}, n_{10y}, n_{15y} and n_{20y}, respectively.

3.2.4.2. Selection of subsets for NIR models

Subsets meeting these sizes (n_{5y}, n_{10y}, n_{15y} and n_{20y}) were selected to build NIR models (see next section for further details). Regardless of its size, each subset was composed by a selection of spectrally representative samples. For that, a principal component analysis (PCA) was performed with the spectra (pre-processed with first derivative, Savitzsky-Golay); then, the scores over the first and second PC were subjected to cluster analysis, and the most central samples of clusters were selected. The subset of size n_{20y} contains the samples previously selected for the subset with size n_{15y}. The subset of size n_{15y} contains the samples previously selected for the subset with size n_{10y}, and the subset of size n_{10y} contains the samples previously selected for the subset with size n_{5y}.

3.2.4.3. NIR models

These subsets were used to build NIR models in two different approaches: i) as calibration sets to build geographically local models, and (ii) as spiking subsets employed to adapt a national model. In all cases, the models were calibrated with partial least squares regression (PLSR).

For the first approach (i), one geographically local model was calibrated with each of these subsets. Thus, the size of the calibration set of these four models was equal to n_{5y}, n_{10y}, n_{15y} and n_{20y}. For the second approach (ii), these subsets selected were used to adapt a national model. For that adaptation, we used spiking with extra-weighting (Guerrero et al., 2016), although a new modification was introduced, inspired by the recent findings of the study of Lobsey et al. (2017). For each subset, the main steps were: 1º) the national model (n=3606) was spiked with the subset; 2º) the spiking subset was extra-weighted (following details in Guerrero et al. 2016) and a model is calibrated; 3º) the national samples identified as outliers in the cross-validation are deleted, and a new model is calibrated; 4º) if new samples appear as outliers, a new "clean outliers and re-calibration" cycle is repeated (i.e., the point 3 is repeated). This sequence stops when no national samples are displayed as new outliers or until all the national samples from the general model have been cleared. Consequently, a number of models is generated during these cycles. The final value of SOC assigned to each predicted sample was the averaged value of the predictions obtained with all the models, excluding those produced by the first five and the last five models.

3.2.4.4. Predictions of SOC

The SOC contents in all the collected samples were predicted with the different models. However, only those samples not included in the subsets was used to compare the accuracy of both approaches (geographically local models *vs* national model adapted). In this way, a fair comparison of results is allowed since the characteristics of the prediction set do not vary, being fully comparable. For the analysis of the accuracy, the next prediction performance parameters were obtained: R^2 (determination coefficient), standard error of prediction (SEP), bias, root mean square error of prediction (RMSEP) and ratio of performance to interquartile range (RPIQ). In addition, the SOC values predicted with the different models were used to estimate μ_{NT20} and μ_{CON}, allowing the computation of ΔSOC20 and σ^2_{0-20y} when NIR is used to predict the values of such parameters, and also their deviation respect to the actual values (measured with WB). These deviations (or errors) were denoted as dΔSOC and dσ^2. Different values of these errors have been obtained depending on the approach (geographically local model or adapted national) and on the size of n. In order to evaluate the best model to predict SOC for monitoring purpose, we decided to select the model that better reproduce the mean value of the two compared populations and the pooled variance. These parameters summarize the SOC distribution within each group. The minimum difference that can be detected is influenced by their values (see the following paragraph). So, we chose as the best one the model with the error BIAS close to be null, and with the lower SEP.

3.2.5. Data analysis – Experimental setup

In order to evaluate for what scenario the advantage by NIR (inexpensive increase of the sample size through the use of NIR predictions) can compensate its drawbacks (lower precision and accuracy than the reference method), the minimum detectable difference (MDD) reached by the reference method was compared with MDD by NIR, taking into account the error of the prediction. In particular, for each scenario NIR could be used for OC monitoring, and a cost-efficient benefit was assured when: i) the MDD by NIR predictions is lower than the ΔSOC that has to be detected, and ii) the costs for the analytical data used for the model (as calibration dataset of local models or spiking subset of general models) plus the costs for NIR spectra acquisitions, is lower than the cost of the reference method.

3.2.5.1. MDD

The MDD was computed with the formula described in Equation 3.3, where the MDD is related with the variance and sample size for an arbitrary value of significance level (α) and power (1-β), which are parameters related with the Type I error (the probability of rejecting the null hypothesis when it is true) and Type II error error (the probability of accepting the null hypothesis when it is false), respectively.

$$\mathrm{MDD} = \sqrt{\frac{(t_{(\alpha/2)}+t_{(1-\beta)})^2 \times 2(\sigma^2)}{n}} \qquad \text{[Eq. 3.3]}$$

being t(α/2) is 1.96 (at α=0.05), t(1-β) is 1.282 (β=0.90), and σ^2 is the sample variance.

For this study the MDD is defined as the smallest difference that can be detected between organic Carbon contents.

For a given condition (power, variance, etc.), the Eq. 3.3 allows to know the MDD that can be assessed with an arbitrary sample size. However, when the method presents error (such as in NIR), the MDD is higher for that given sample size, due to the poorer capacity respect to an errorless method (such as WB). In the following paragraphs we reported the method that we adopted to quantify how much the error affect the MDD when an establish number of samples was used.

3.2.5.1.1 MDD with reference method (Walkley-Black)

We computed the MDD obtained with the different subsets (n_{5y}, n_{10y}, n_{15y} and n_{20y}) in each scenario. The σ^2 used in each scenario were those described in section 2.3. Since the SOC in these samples was measured with WB, these MDD values were referred as MDD_{WB}.

3.2.5.1.2. MDD using NIR

We computed the MDD obtained with 148 samples, which is the sample size reached once a spectroscopic model is available to predict the SOC with NIR in all samples. These MDD values were referred as MDD_{NIR}. However, the simple computation of the MDD_{NIR} using Eq. 3.3 and n=148 implies the assumption that NIR predictions are error-free measurements, which is ideal but unrealistic. Consequently, it assumes that the parameters needed to compute the MDD in Eq. 3.3 are described by the NIR predictions without any error. Therefore, with the aim to obtain a more realistic value of MDD_{NIR}, we computed a corrected MDD_{NIRc}, using Eq. 3.4,

$$MDD_{NIRc} = \left(\sqrt{\frac{(t_{(\alpha/2)}+t_{(1-\beta)})^2 \times 2(\sigma c^2)}{n}}\right) + d\Delta SOC \qquad \text{[Eq. 3.4]}$$

where the errors from NIR predictions are considered and they exert a negative influence the MDD. For that, the two errors obtained in section 2.4.4 were considered to produce a corrected MDD:

i) the error of NIR predictions to provide a true value of the σ^2, which was computed in section 2.4.4, was used to compute the σc^2 using Eq. 3.5:

$$\sigma c^2 = \sigma^2 + d\sigma^2 \qquad \text{[Eq. 3.5]}$$

The deviation respect to the true value ($d\sigma^2$), obtained with the reference method in the 148 samples (section 2.4.4), could be positive or negative. It was assumed the worst hypothesis according to which the variance estimated from the model was higher of the real one and the difference between mean was lower, Indeed, a negative value would decrease the MDD_{NIRc} respect to the MDD_{NIR}, providing an apparent (false) improvement. Therefore, this deviation was always used with positive sign, leading to a higher σc^2, consequently increasing the MDD_{NIRc} respect to MDD_{NIR}. In that way the larger the error, the higher the penalization in the MDD_{NIRc}.

ii) the error of NIR predictions to provide the true ΔSOC (dΔSOC). This error was directly added (with positive sign) in Eq. 3.4, penalizing the MDD_{NIRc} (i.e., increasing the MDD_{NIRc}).

3.2.5.2 Comparison of costs of WB and NIR at different scenarios

For a given conditions (power, variance, etc.), the Eq. 3.3 allows to know the number of samples needed for detecting a given difference. Due to the poorer capacity respect to an errorless method, a method with errors would need a larger sample size to achieve the same MDD. Indeed, the increase in the sample size decreases the MDD. For instance, WB is not really errorless when compared to dry combustion (De Vos et al., 2007), anyway it may be assumed as an "errorless method" since it is accepted as a reference methods. On the other side, NIR can be assumed as a method with errors since its use is generally based on calibrating a multivariate model on reference data and models alway add an error in prediction, compared with the refence data. The Eq. 3.4 allowed to know the number of samples required to detect a given difference, when the method affected by error was adopted. Therefore, we can estimate what is the additional sample size needed to achieve the requested MDD. The higher is the error, the higher the required increase in number size. The extra-sample needed for the compensation implies some additional costs, which depend on the required number size and the costs for acquiring spectral signatures. Although NIR is a cheap method, an increase in sample size requires sampling efforts and sample preparation, which implies some costs. So, we have analysed costs at different conditions. We computed the costs needed in traditional investigation by means of WB and they were compared with costs using NIR.

For the estimation of costs using WB, we used the Eq. 3.6:

$$\text{Cost\_WBi (in euros)} = ni \times 10 \qquad \text{[Eq. 3.6]}$$

where 10 is the cost (in euros) for the SOC analysis per sample, and ni is the sample size needed to meet the required power for an arbitrary ΔSOC assessment according with Eq. 3.2.

For the estimation of the cost using NIR, we used the Eq. 3.7, which consists in two parts:

NIR_costi (in euros) = Eq. 3.8 + Eq.3. 9 = cost of reference analyses of local samples + scanning cost [Eq. 3.7]

being:

$$\text{cost of reference analyses of local samples (in euros)} = nss \times 10 \qquad \text{[Eq. 3.8]}$$

where nss is the size of the local dataset. Here, four local dataset have been used in this study (8, 12, 22, 56).

$$\text{scanning cost (in euros)} = nt \times 1 \quad = (ni + nc) \times 1 \qquad \text{[Eq. 3.9]}$$

where nt comprises the ni and nc, being ni is the sample size needed to meet the required power for an arbitrary ΔSOC assessment according with Eq. 3.2, and nc is the extra-sample size needed to compensate.

The nc varies as consequence of the errors, which in turn, may depend on the effort invested in adapt (in general, a model adapted with a small spiking subset is likely to predict with larger errors than another adapted with a larger spiking subset).

In this study we are assuming a cost of 1 euro per sample in Eq. 3.8, resulting in a ratio 10:1 for the WB and NIR per sample. The value assigned to the ratio in this study is merely operative, and it can be changed easily. The precise calculation of its value depends on several factors, such as the reference method (WB, elemental analyser, etc.), sampling and sample support, or even the inclusion of fixed costs (O'Rourke and Holden, 2011; Nocita et al. 2015) and also the management of toxic wastes generated in laboratory; such detailed calculation is out of the scope of the paper. It is also worth to highlight that these prices were including the field work (travel, sampling, etc.), sample preparation, analysis and wastes disposal (in the case of WB). As consequence, the values are probably unrealistic, but they can be easily rescaled.

The comparisons have been carried out in the four scenarios: 5, 10, 15 and 20 years. Since the pooled variance was different in each scenario (after 5, 10, 15 and 20 years), the comparisons were carried out separately.

3.3. Results and discussion

3.3.1. SOC contents under conventional tillage, 20 years of no tillage, and in the hypothesised scenarios

SOC values under conventional and no-tillage conditions are graphically synthesized in the figure 3.2. The populations under conventional tillage and the population under 20 years of no tillage are significantly separated according to the mean and standard variation values. The SOC contents were ranging from 0.75 to 1.55 %SOC in CON and from 1.02 to 2.66 %SOC in NT20. The mean values in CON (μ_{CON}) and NT20 (μ_{NT20}) were 1.13 and 1.67 %SOC respectively. Therefore, the SOC variation after 20 years of the land management change (ΔSOC20) was 0.54 %SOC, corresponding at the 48% of the original value (CON).

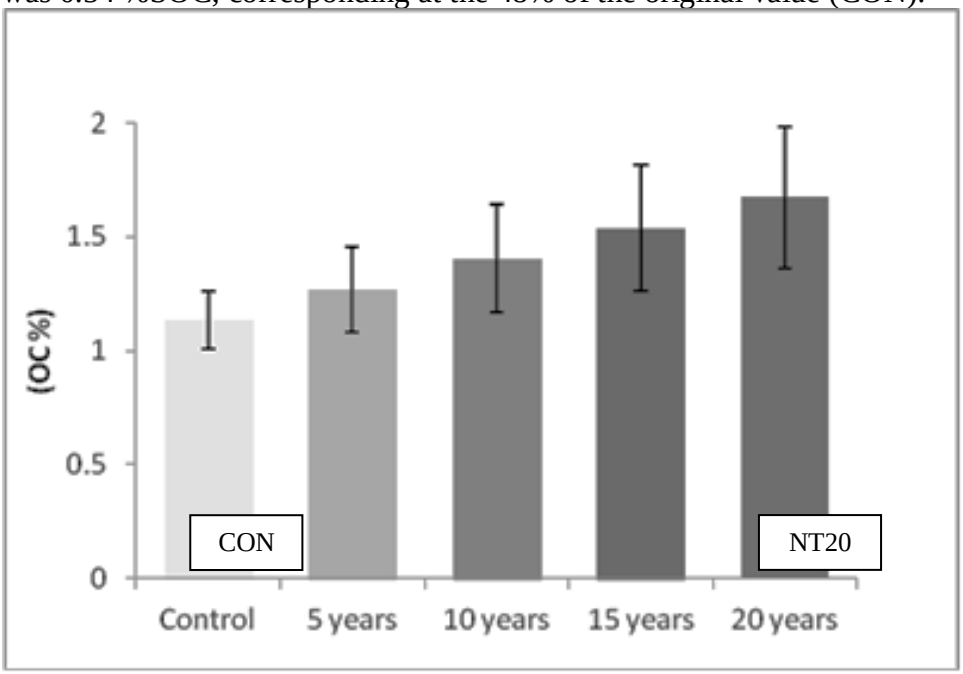


Figure 3.2: Mean values (± standard deviation) of Organic Carbon in the conventional field, after 20 years of no tillage practice, and in the scenarios after 5, 10 and 15 years no tillage.

Based on our linearity hypothesis, the mean increased of 0.14% after 5 years (from 1.13 g kg^{-1} to 1.27, corresponding at the 12% of the original value (CON).), of 0.27% after 10 years

(from 1.13 g kg^{-1} to 1.40, corresponding at the 24% of the original value (CON)), and of 0.41 % after 15 years (1.13 g kg^{-1} to 1.54, corresponding at the 36% of the original value (CON)) (Fig. 2). Janzen et al. (1998) suggested that gains in SOC upon adoption of improved practices will presumably follow an asymptotic curve, likewise the loss of SOC upon adoption of practices with lower C-retention. The temporal change in SOC due to tillage practices is supposed to vary with soil texture, climate and biomass return (Dıaz-Zorita and Grove, 2002). The duration of SOC gain may be very different Campbell et al. (1995) estimated a duration of 70 to 80 years, with most of the change occurring in 5 years, while Izaurralde et al., (1997) hypothesized a potential gain of 5 decades long. A rapid SOC accumulation was observed initially with adoption of no tillage followed by minimal changes after 5–10 years by Staley et al. (1988) and Dick et al. (1991). However, based on a comparison of soils at several sites under NT, there was no clear trend in tillage-induced changes in SOC as a function of time in a report by Paustian et al. (1997). Even if a liner increase was here adopted for creating the scenarios, our results about the effect of land management change on SOC were consistent with the outcomes of previous studies. As early as the end of the 1990s, Garten and Wullschleger (1999) worked in bioenergetics crops of southeastern USA and they found an increase of ≈10 to 15% of existing SOC after 5 years from changing in land use; Angers et al. (1998), found an increase of about 10% in SOC after 11 years from the introduction of no-tillage practice in the eastern Canada.

For a general overview of the question, it has also to be taken into account that in this study, it was considered the surface soil (10 cm depth) and the observed surface SOC concentration change upon 20 years of no-till may not reflect the actual SOC concentration change in the whole soil profile. Indeed, the current scientific knowledge on the effect of tillage on SOC stocks shows no effect of this management practice on SOC stocks in the whole soil profile (Virto et al., 2012).

As regards the variance, it also was higher under no-tillage condition: The variance in NT20 ($\sigma^2_{NT20} = 0.096$) was about 6 times bigger than the observed in CON $\sigma^2_{CON} = 0.016$). Based on our linearity hypothesis, the variance increased from 0.016 to 0.036 after 5 years, from 0.016 to 0.056 after 10 years, and from 0.016 to 0.76 after 15 years.

.

3.3.2. Sample size to detect ΔSOC in the scenarios

The sample size needed to detect such ΔSOC in each temporal scenario was obtained using a power analysis. The number of samples per group (per field) needed to detect ΔSOC_5, ΔSOC_{10}, ΔSOC_{15} and ΔSOC_{20} was 28, 11, 6 and 4 respectively (table 3.1). When the change is cumulative the change to detect (ΔSOC) increases, and consequently the number of samples needed decreases.

As the variance is not constant, but it had an increase along the time, after the introduction of no tillage, and *n* is a function both of the MDD and of the variance, we reported a different function for each scenario (fig. 3.3).

Looking at previous studies over the amount of data needed for detecting SOC increase, authors reported a quite wide range of value. For instance, Garten and Wullschleger (1999) suggested a sample sizes of 16 samples for detecting a 10-15% of organic carbon increase occurring after 5 years ($1 - \beta = 0.90$). Kucharik, et al. (2003) concluded that 40 to 65 paired sites needed to achieve an 80% confidence level ($\alpha = 0.05$; $\beta = 0.20$) in soil C rates after at least 8 years from planting for the Conservation Reserve Programs of Wisconsin.

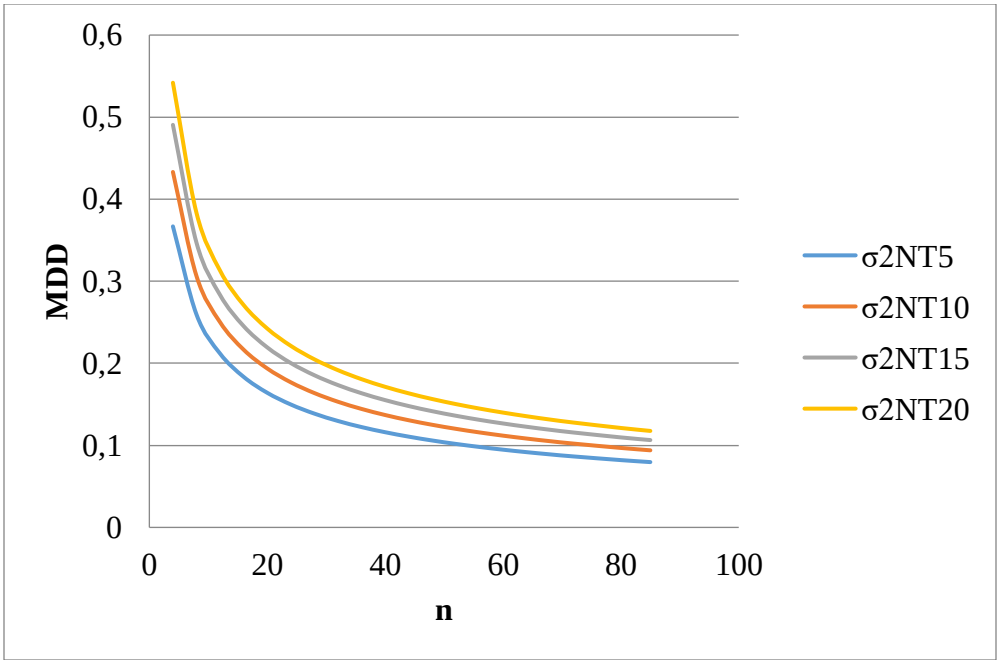


Figure 3.3: Curves of the MDD as function of the sample size (*n*) and the pooled variance (σ2NT5: pooled variance after 5 years of no tillage, etc.)

Table 3.1: Main statistics synthetic parameters of OC after 20 years of no tillage and in the hypothetic scenarios based on a linear increase during 20 years of no tillage; ΔSOC values was calculated according to Eq.3.1; *n* was calculated by Eq.3.2.

	Condition in the scenario		Parameters of Eq. 3.2		
Years from NT introduction	μ_{NT}	σ^2_{NT}	ΔSOC	σ^2_{0-y}	n
5	1,27	0,036	0.14	0,0256	56 (28 per field)
10	1,40	0,056	0.27	0,0357	22 (11 per field)
15	1,54	0,076	0.41	0,0458	12 (6 per field)
20	1,67	0,096	0.54	0,0559	8 (4 per field)

3.3.3. NIR models and predictions

The four geographically local models were calibrated with 56 (28×2), 22 (11×2), 12 (6×2) and 8 (4×2) samples. These numbers correspond with the sample size needed for detect the ΔSOC5, ΔSOC10, ΔSOC15 and ΔSOC20, according with the power analysis (at α=0.05 and 1-β=0.90). These four models were used to predict the SOC in the prediction set. These subsets were also used to spike a national scale model.

In our opinion, some NIR errors are more important than others. The RMSEP is a measure of the accuracy, and it is composed by the precision and the bias. None of these parameters are directly useful to inform about the capacity of the predictions to describe the ΔSOC and variance. The bias in each group is directly related with the assessment of the ΔSOC. So, the predictions with a small or null bias (in each group) should be the most interesting property of a model. Values of bias close to zero are frequently reported, and spiking with EW is one outstanding technique for that purpose. Under that circumstance, bias close to zero, the RMSEP is due to SEP, being both irrelevant. The capacity to describe the variances is less clearly associated with errors, and probably related with the R^2.

The main results are shown in Table 3.2.

As expected, in the four geographically local models, the number of samples used to calibrate the models had a direct positive influence on the quality of predictions.

Unspiked models showed a strong improvement in terms of RMSEP if compared with the local models, due to the vert bigger size of the dataset. However, the predictions obtained were clearly biased and RPD in very low. Very similar results were obtained for the spiked national scale models: the predictions obtained were clearly biased. This effect has been observed by other authors (Guerrero et al. 2016), and it is probably due to the large size of the national model and the small size of the spiking subsets. Even with the largest sized spiking subset (56 samples), this effect was small (Figure 3.4). An improvement in prediction accuracy was observed when the spiking subsets were extra-weighted, and models were calibrated after successive removal of outliers. The most relevant improvement respect to the spiked models is the reduction of the bias, although some improvements in the other indexes were pointed out.

As it was observed in the geographically local models, the increase in the number of samples used had a positive influence in the quality of the predictions. However, it is clear that the benefits obtained by the analysis of 56 samples, which is a very important analytical effort, are not worth in comparison with moderate efforts.

Regardless of the number of samples considered, the quality of the predictions was systematically lower when geographically local models were used. Therefore, the predictions obtained with that approach were not considered for the next analysis. As a whole extra-weighted spiked calibration outperformed the geographically-local one. We chose the extra-weighted spiked calibrations as the best ones, since they were the less affected by BIAS likewise the local ones. but they presented better slope and offset values (fig. 3.4). It means that the BIAS was distributed equally over all the data, so that the difference between the two population should me better estimated. The extra-weighted spiked calibrations also presented a better precision and correlation, particularly when only 12 samples were used in the models. In the table 3.3 it was reported the statistics values of mean and variance that occurred when the extra-weighted spiked calibrations were carried out with the sized spiking subsets of 56, 22, 12 and 8 local samples, and the differences with the real data (based on 74+74 WB samples) were also reported.

Table 3.2: Prediction performance parameters for local, initial general calibration (Unspiked), spiked (Spk) and extra-weighted spiked calibrations (spkEW cleared): RMSEP (root mean square error of prediction), SEP (standard error of prediction), BIAS, R^2 (coefficient of determination), and RPD (ratio of performance to deviance). Predictive parameters were estimated on the same validation dataset of 96 samples.

predictive parameter	*n*	**Calibrations**			
		local	**Unspiked**	**Spk**	**spkEW cleared**
RMSEP	*0*	-	0.72	-	-
	8	0.20	-	0.69	0.19
	12	0.18	-	0.68	0.15
	22	0.17	-	0.65	0.14
	56	0.16	-	0.57	0.13
BIAS	*0*	-	0.60	-	-
	8	0.01	-	0.57	0.03
	12	-0.04	-	0.55	-0.02
	22	-0.03	-	0.52	-0.02
	56	-0.01	-	0.44	0.00
SEP	*0*	-	0.40	-	-
	8	0.21	-	0.40	0.19
	12	0.18	-	0.39	0.14
	22	0.16	-	0.39	0.14
	56	0.16	-	0.37	0.13
R^2	*0*	-	0.87	-	-
	8	0.74	-	0.87	0.84
	12	0.79	-	0.87	0.86
	22	0.85	-	0.87	0.86
	56	0.84	-	0.87	0.88
RPD	*0*	-	0.52	-	-
	8	1.82	-	0.54	1.98
	12	2.06	-	0.55	2.56
	22	2.24	-	0.57	2.59
	56	2.38	-	0.65	2.89

Table 3.3: Mean and variance obtained by 148 predicted SOC values when the extra-weighted spiked calibrations were carried out using spiking subsets of 56, 22, 12 and 8 local samples, and their differences with the real data ($d\sigma^2$: deviation from the real pooled variance; $d\Delta SOC$ deviation from the real increase in mean SOC) The real data were calculated on the base of the 74 samples from each fields analyzed by WB.

	spkEW cleared models				Real
Statistical parameters	NIR56	NIR22	NIR12	NIR8	data
N of samples	74+74	74+74	74+74	74+74	74+74
μ_{CON}	1.131	1.116	1.106	1.133	1.130
μ_{NT}	1.671	1.655	1.666	1.750	1.673
ΔSOC (μ_{NT} - μ_{CON})	0.540	0.539	0.560	0.617	0.543
$d\Delta SOC$	*0.003*	*0.005*	*0.018*	*0.074*	--
σ^2_{CON}	0.0090	0.0161	0.0150	0.0122	0.0155
σ^2_{NT}	0.0791	0.0984	0.1110	0.1662	0.096
Pooled σ^2 [$1/2(\sigma^2_{CON}+\sigma^2_{NT})$]	0.0440	0.0573	0.0630	0.0892	0.0559
$d\sigma^2$	*0.0118*	*0.0014*	*0.0071*	*0.0333*	--

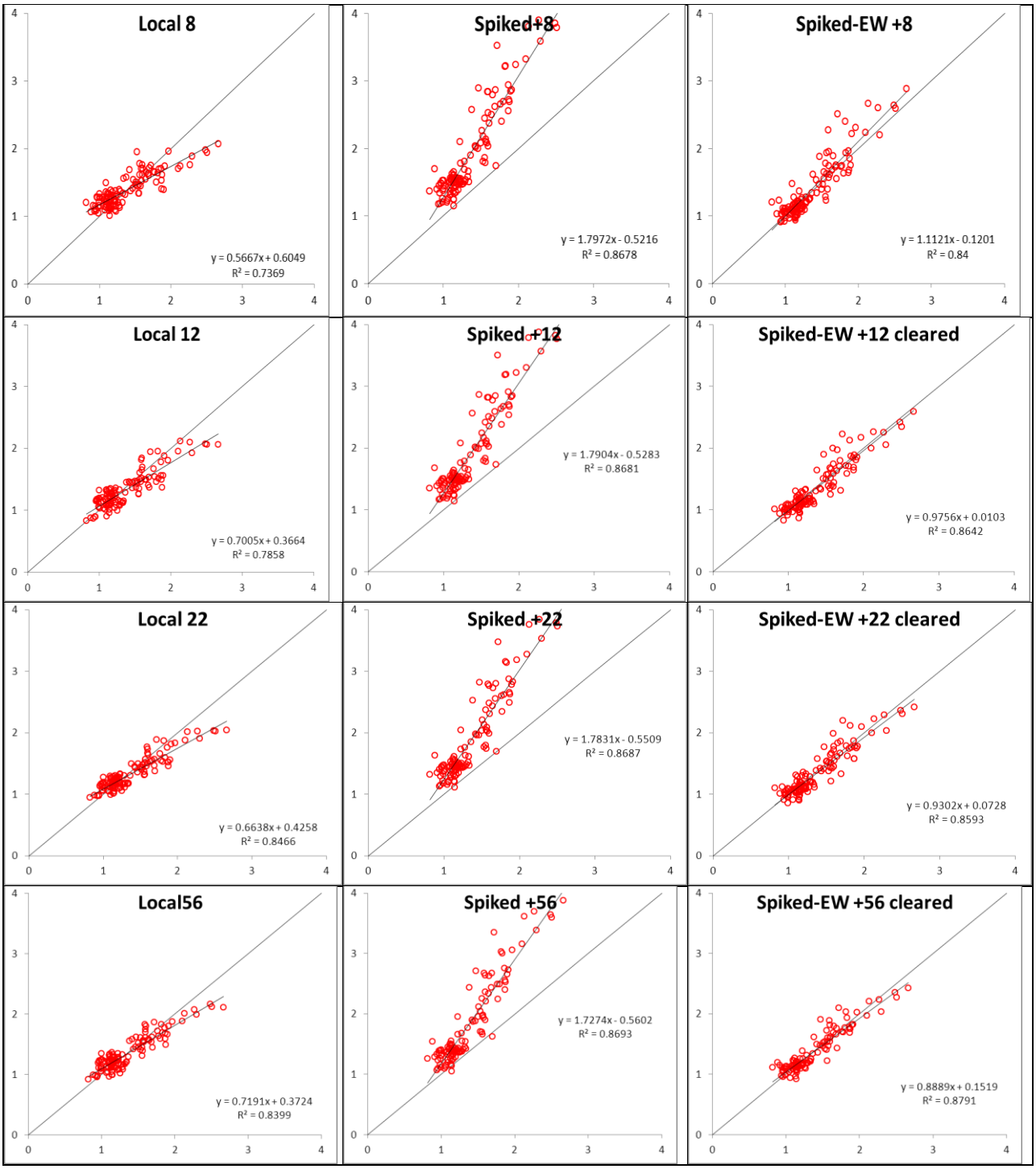


Figure 3.4: Representative illustration of observed vs predicted OC%. Left: predictions obtained from geographically-local models. Centre: predictions obtained from spiked models. Right: predictions obtained from models with extra-weighted spiking dataset. First row: models obtained with a local sample size=8 (4+4). Second line: models obtained with a local sample size=12 (6+6). Third line: models obtained with a local sample size=22 (11+11). Fourth line: models obtained with a local sample size=56(28+28). Axes X: observed OC%; axes Y: predicted OC%.

3.3.4. MDD_{WB} vs MDD_{NIR}, y MDD_{NIR} vs MDD_{NIRc}

After five years of the change, the sample size needed to detect ΔSOC_5 is 28 data WB per group. The MDD_{WB} needs 28 samples per field. According with the linear cumulative change assumption, after five years the ΔSOC_5 should be ~0.14, and the σ^2_{0-5y} should be around 0.025. When a NIR model is used, the sample size was increased up to 148 (74 per group). In theory, the MDD that can be reached by means of NIR (MDD_{NIR}), when the pooled variance is σ^2_{0-5y},and independently from the spiking dataset size adopted model, should be 0.085, a clearly lower value.

After ten years of the change, the number of samples needed to detect the ΔSOC_{10} (~0.27), with the desired power is 22 (11 per group), while, when the pooled variance is $\sigma^2_{0\text{-}10y}$, the MDD_{NIR} should be 0.101.

After fifteen years of the change, the number of samples needed to detect the ΔSOC_{15} (~0.40), with the desired power is 12 (6 per group), while, when the pooled variance is $\sigma^2_{0\text{-}15y}$, the MDD_{NIR} should be 0.114.

Finally, after twenty years of the change, the number of samples needed to detect the ΔSOC_{20} (0.54), with the desired power is 8 (4 per group), while, when the pooled variance is $\sigma^2_{0\text{-}20y}$, the MDD_{NIR} should be 0.126.

In all the four cases, the MDD_{NIR} was a clearly lower value than the ΔSOC to detect. However, these values are true only when the NIR predictions are error-free measurements. Therefore, MDD_{NIR} shouldn't be considered as the realistic because predictions contain errors. In order to provide a more realistic value of the MDD_{NIR}, we have used the errors to produce a corrected value (MDD_{NIRc}).

The first error to consider was related with the variance.

As first, we took into account what could happened monitoring the SOC increase after 5 years by means of the NIR model based on a spiking dataset of 56 samples.

We computed the deviation from the real pooled variance ($d\sigma^2$) and the deviation from the real increase in mean SOC ($d\Delta SOC$) (table 3.3). There was an absolute difference in the variance (0.0118) respect the "true" value. The higher the lack in the accuracy, the poorer the capacity to quantify the ΔSOC, and therefore the higher the MDD_{NIRc}. So, this inaccuracy was added to increase the variance, in the equation 3.5: $d\sigma^2$ was added to the variance attended after 5 years. Consequently, MDD_{NIRc} will result higher than MDD_{NIR}. The second step to correct the MDD_{NIR} is related with the capacity of NIR prediction of reproducing the variation ΔSOC with accuracy. The deviation respect to the true value was 0.003. As we did with the variance, any deviation respect to the true value should be understand as a deterioration in the MDD_{NIRc} respect MDD_{NIR}. Therefore, this deviation (its absolute value) was added to increase the MDD_{NIRc} respect to the MDD_{NIR}. Once corrected, the MDD_{NIRc} was 0.106, which is still lower than the MDD_{WB} (0.14). Therefore, there is an advantage when NIR is used.

When the national model was adapted with a spiking subset composed by 22 samples in the same scenario of 5 years of notillage, the MDD_{NIRc} was slightly different from the previous one (the capacity of predictions to describe the variances and ΔSOC is different from the model with a spiking subset composed by 56 samples) . However, the analytical effort needed was clearly smaller, and the difference of MDD_{NIRc} (0.092) respect to MDD_{WB} (0.14) is substantial. When 12 samples were used for the adaptation of the NIR model, the quality of the predictions

was still high, allowing to describe accurately the variances and ΔSOC, and therefore leading to a MDD_{NIRc} (0. 0.113), lower than the MDD_{WB} (0.14). When only 8 samples were used as spiking subset, the MDD_{NIRc} (0.203) was clearly higher than the MDD_{NIR} (0.085) due to the higher errors in the predictions. Cosequently , differently from the spiking datasets of 56, 22 and 12 samples, the MDD_{NIRc} was no lower than the MDD_{WB} (0.14) and it cannot be used to predict the difference expected after 5 years.

The results were also suggesting that the size of the spiking subset should be optimized, since a large size implies an effort in analysis which seems unnecessary in terms of provide a lower MDD. Similarly, a spiking subset of very small size might be insufficient to adapt properly the model to the local conditions, and poor quality predictions may ruin the advantages of NIR (as a large sample size).

After ten years, the ΔSOC10 is 0.27, and the variance σ^2_{0-10y} should be around 0.035. Therefore, a minimum of 11 samples per group are needed to meet with the desired power. When 22 (11+11) samples were used as spiking subset, the predictions allowed to obtain a MDD_{NIRc} (0.107) lower than the ΔSOC10. When the number of samples analysed by the reference method is below 22, the MDD_{WB} is higher than the ΔSOC10. Conversely, predictions obtained with models adapted with spiking subsets of sizes 12 (6+6) or 8 (4+4) were providing values of MDD_{NIRc} below the ΔSOC10 (0.214 and 0.217 respectively).

After 15 years and 20 years, the observations are very similar. The number of samples needed per group with WB approach is 6 and the ΔSOC15 is equal to 0.41. The MDD_{NIRc} was clearly lower than ΔSOC15, both when the same number of samples (12=6+6) were used for spiking the NIR model (MDD_{NIRc} 0.140) and when the number of spiking samples was lower (8=4+4) (MDD_{NIRc} 0.223).

After 20 years, the ΔSOC20, that has to be checked, is 0.54, and σ^2_{0-20y} is 0.055; therefore only 4 samples per group are needed to meet with the pre-established power. Despite of the lower accuracy of models adapted with 8 samples as spiking subsets, the MDD_{NIRc} (0.233) was still lower than the ΔSOC20. So, these results were indicating that, for equal effort made to analyse samples with the reference method, NIR was always providing a further advantage respect to traditional approach: when the analysed samples were used to adapt a national model a lower MDD can be achieved by means of NIR, in comparison with MDD_{WB}. (fig. 3.5)

It seems that the benefit of NIR trends to diminish when the ΔSOC increases. So, waiting until a large ΔSOC value can be measured with a relatively low number of samples seems to be an option to avoid important costs in monitoring. Indeed, NIR grows its advantage in a frequently observed option. However, the reliability of the results could be low for small sample size. Can we expect a robust result when only four samples per group are compared? Can we provide a credible measure of the ΔSOC with this small sample size? For this reason, we have also evaluated the convenience to extend the sample size above the strictly needed number to meet with the power. For that, we have analysed different sample sizes.

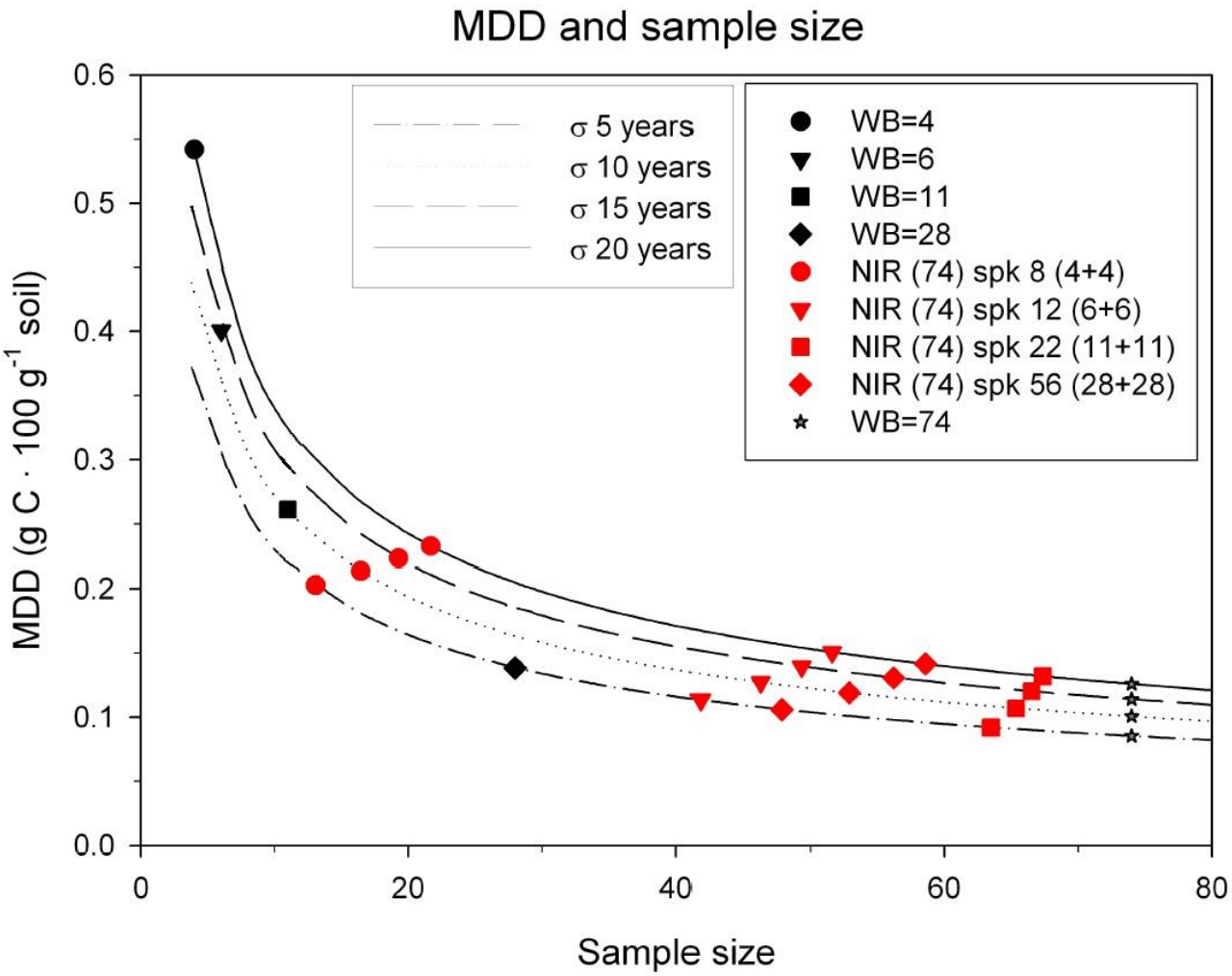


Figure 3.5: Curves of the MDD as function of the sample size (n) and the mean variance of the four scenarios after 5,10, 15 and 20 years of no-tillage. Black symbols point at the MDD_{WB} that was for each scenario. On each line there are a black symbol and 4 red ones. The res symbols represent the MD_{DNIRc} that can be obtained from data estimated by NIR models with different number size of the spiking dataset. The stars stood for the MDD reached when all the 148 (74+74) WB data were used. It coincides with the MDD_{NIR} before the correction.

3.3.5. Estimation of economic efforts and comparison of costs of WB and NIR at different scenarios

For the estimation of economic efforts, we chose to adopt the proportion between VisNIR diffuse reflectance spectroscopy and laboratory analyses methods, reported by O'Rourke and Holden (2011). They calculated the costs per sample, analytical accuracy, and time involved in SOC analysis. They found that Vis-NIR spectroscopy outperformed laboratory analyses for the lower price (10 times cheaper). In the figure 3.6, it was reported the comparison of the costs of both approaches (DMD_{WB} and DMD_{NIRc}), with the aim to see if such compensation is convenient in terms of cost benefit, in the different scenarios (namely at different variance and ΔSOC to detect). Although NIR is cheaper than WB, NIR has costs that we need to consider and may

restrict its applicability. Thus, for NIR, we took into account the costs for analyses with the reference method of the spiking subsets plus the cost for scanning the samples, which is10 times smaller, but it can result in an important cost if thousands of samples are needed for errors' compensation.

Each graph of figure 3.6 referred to a scenario of our study, that is, to a different variance. The red line denotes the change to be detected in each scenario with the pre-established power using the WB. The grey area refers to the cost by WB for the analysis of the required samples according to the Eq. 3.3 and lines denote the NIR costs when different spiking datasets were used.

In the left side of the graphs, costs increased very quickly at the decreasing, as the increase of required samples when the ΔSOC to be detected decreases, follows an exponential trend (Eq. 3.3, eq. 3.4) For NIR, this increase of costs was faster and it started at different values of ΔSOC according to the different accuracy of the NIR models. It was due to the fact that when the ΔSOC to be detected approaches to the error measuring the difference of means, the resulting MDD_{NIRc} requires a very large number of samples. Therefore, the higher is the error, the bigger is the difference to detect at which rapidly NIR advantages decreases.

Looking at the scenario of 20years of NT and at the NIR model spiked with 8 samples, for values of ΔSOC larger than 0.55, the WB is cheaper because we only need to analyse a few samples in the laboratory, which may be even a smaller number than the samples forming the spiking subset. For ΔSOC smaller than 0.1, the sample size needed to compensate the errors rapidly increases, and therefore the overall costs are larger. Since the model spiked with only 8 samples was the most affected by error, it is also the which one that firstly loses its cost benefit for little ΔSOC.

Looking at the model spiked with 12 samples (6+6), as the spiking subset is a little bigger that the previous one, the initial cost is slightly higher. However, this spiking subset allowed a successful adaptation of the model, and the errors were smaller. As consequence, the compensation effort is not too big, and the massive rise in costs starts at very lower value than the previous model. The use of this NIR model becomes a cheaper approach respect to WB for ΔSOC values below 0.45.

For the model spiked with 22 samples (11+11), the analysis of a large spiking subset generates high initial cost As consequence, it only becomes cheaper than WB for ΔSOC values below 0.3. As in the previous case, a large spiking subset usually guarantees accurate predictions, and this positive effect implies a small compensation effect. Consequence, at very little ΔSOC (around 0.1), the costs equal to those obtained with a smaller spiking subset (6+6), and at smaller ΔSOC, NIR model spiked with 22 samples outperforms the which one spiked with 12 samples in terms of saving money. Thus, it seems worth to extend the size of the spiking subset if small changes are expected. However, the results found with the largest spiking subset (56=28+28) suggest that the benefits of such increase are limited. Indeed, it only becomes cheaper than WB when ΔSOC is lower that around 0.2 and is more expensive respect to a spiking subset composed by 12 (6+6), when ΔSOC is higher that around 0.05.

Very similar trends were obtained among for the models in the other scenarios, namely at other variance values. What mostly change was the value of ΔSOC to detect. Up to 5 years, the difference to detect results so little that models with big errors have to be excluded, while all the other models outperformed WB in terms of cost benefits. When the conditions of this scenario occurred the NIR models with spiked subset of 12 and 22 still outperformed WB, but not the NIR model with spiked subset of 56. At the condition of 10 years of no tillage, the sample size required by WB method was small enough to make WB a little be more convenient than 22 spiked model, while NIR models with spiked subset of 8 and 12 still outperformed WB. After 15 years, the SOC change started to be bigger enough to make NIR not convenient, due to the little sample size required by WB.

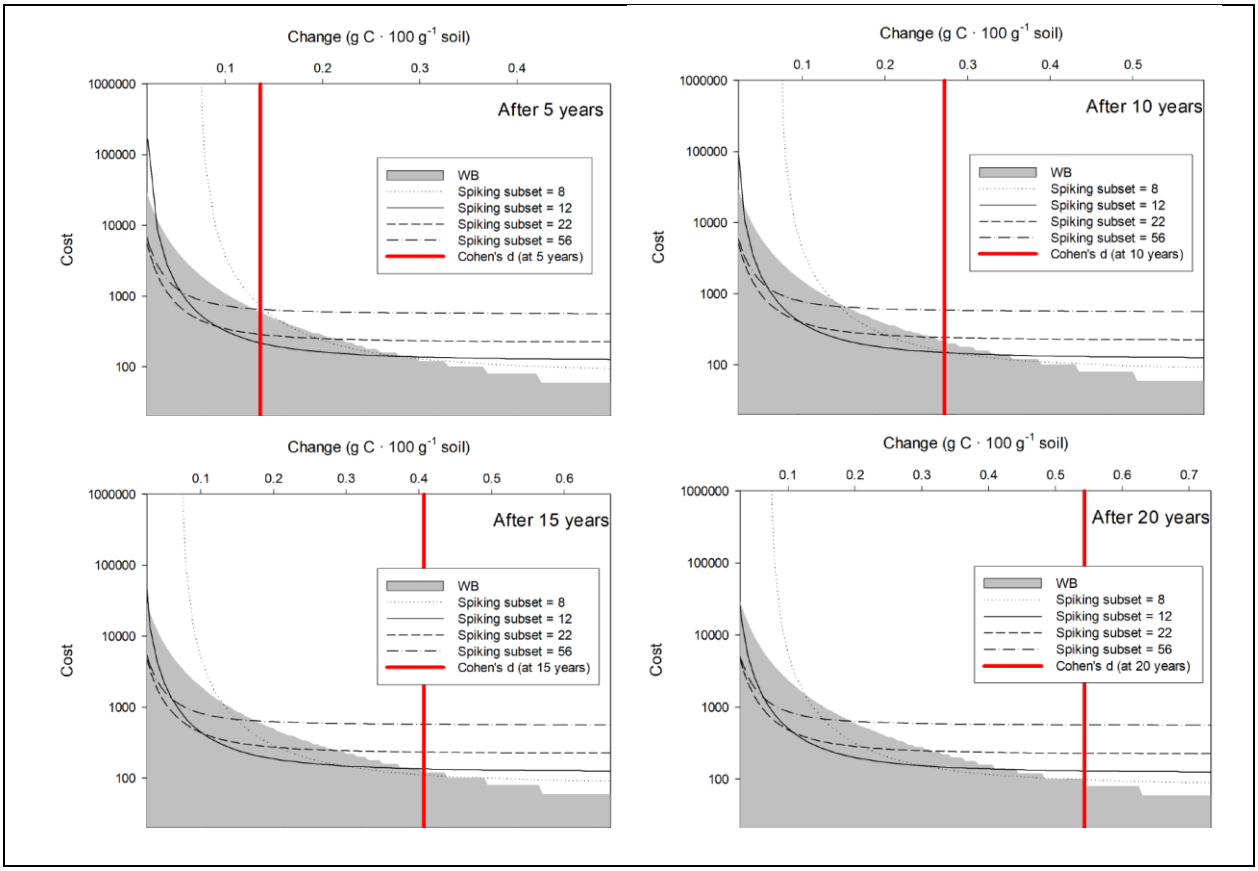


Figure 3.6: comparison of the costs of both approaches (DMD_{WB} and DMD_{NIRc}), in the 4 scenarios. The red line denotes the change to be detected in each. The grey area refers to the cost by WB for the analysis of the required samples and lines denote the NIR costs when different spiking datasets were used.

3.4. Conclusions

Comparing SOC by means of NIR required to select a model with very low BIAS, in order to achieve the lowest error in reproducing real differences between the compared data. Extra-weighted spiked models demonstrated to be the most appropriated for the issue.

NIR approach demonstrated to be a powerful instrument to detect change in SOC content: NIR always outperformed WB in terms of MDD when the same number of analyzed data was employed.

We can assume an economic benefit with NIR spectroscopy, when the number size required to detect C difference by conventional wet chemistry is bigger than what we need for accurate NIR calibration. But, when the C difference to detect is very big or when the populations show a very little pooled variance, a very small number of conventional wet chemistry data was required for detecting differences. In these cases, the cost-effective advantage from NIR spectroscopy may be reduced or even disappear.

So, NIR is a competitive technique when small changes should be detected, as those expected in less than 10 years. However, its advantages are restricted by the size of the spiking subset, which requires an equilibrium: i) a small spiking subset can result in poorly adapted models, which implies a large compensation that might be unpractical (more expensive); (ii) when a large spiking subset is used, the compensation effort is small, but it requires an important initial effort, which also decreases the benefits. A bigger spiking subset should be preferred when very little differences are expected (<0.1%) as those hypothesised for less than 5 years, or when large variance may occur.

Moreover, it is important to consider that if monitoring activity provides repeated assessments, the cost will favour the use of NIR, since there is no need to analyse the spiking subset (i.e., the model does not need to be adapted each time) and those initial efforts invested on adapting the models might be less important in the overall budget. Thus, after the repetition of several assessments (successive events), the NIR is expected to surpass WB. Further studies may focus on a quantification of this aspect for a very exhaustive discussion over this topic.

Acknowledge

This study started during my period of doctoral study at the University Miguel Hernandez de Elche Department of Agronomica y Medio Ambiente. I want to thanks very much the Tutor of the host University, the professor Cesar Guerrero, without which this study would not have been realized.

1.1.1.4 References

Andrews, S. S., Karlen, D. L., and Cambardella, C. A., 2004. The soil management assessment framework: A quantitative soil quality evaluation method. Soil Sci. Soc. Am. J. 68, 1945–1962.

Angers, D. A., Bolinder, M. A., Carter, M. R., Gregorich, E. G., Drury, C. F., Liang, B. C., ... & Martel, J., 1997. Impact of tillage practices on organic carbon and nitrogen storage in cool, humid soils of eastern Canada. Soil and Tillage Research, 41(3-4), 191-201.

Campbell, C.A., McConkey, B.G., Zentner, R.P., Dyck, F.B., Selles, F., Custin, D., 1995. Carbon sequestration in a Brown Chernozem as affected by tillage and rotation. Can. J. Soil Sci. 75, 449±458.

Cécillon, L., Barthès, B. G., Gomez, C., Ertlen, D., Génot, V., Hedde, M., ... & Brun, J. J. 2009. Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). European Journal of Soil Science, 60(5), 770-784

De Vos, B., Lettens, S., Muys, B., & Deckers, J. A. (2007). Walkley–Black analysis of forest soil organic carbon: recovery, limitations and uncertainty. Soil Use and Management, 23(3), 221-229.

Dıaz-Zorita, M., Grove, J. H., 2002. Duration of tillage management affects carbon and phosphorus stratification in phosphatic Paleudalfs. Soil and Tillage Research, 66(2), 165-174.

Dick, W.A., Edwards, W.M., McCoy, E.L., 1997. Continuous application of no-tillage to Ohio soils: changes in crop yields and organic matter-related soil properties. In: Paul, E.A., Paustian, K., Elliott, E.T., Cole, C.V. (Eds.), Soil Organic Matter in Temperate Agroecosystems. CRC Press, Boca Raton, FL, USA, pp. 171–182.

England, J. R., Rossel, R. A. V., 2018. Proximal sensing for soil carbon accounting. SOIL Discuss., https://doi.org/10.5194/soil-2017-36. In press.

Garten, C. T. Jr. , Wullschleger, S. D., 1999. Soil carbon inventories under a bioenergy crop (Switchgrass): Measurement limitations. J. Environ. Qual. 28: 1 359–1 365.

Guerrero, C., Wetterlind, J., Stenberg, B., Mouazen, A. M., Gabarrón-Galeote, M. A., Ruiz-Sinoga, J. D., ... & Rossel, R. A. V., 2016. Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy?. Soil and Tillage Research, 155, 501-509.

Kucharik, C. J., Roth, J. A. and Nabielski, R. T., 2003. Statistical assessment of a paired-site approach for verification of carbon and nitrogen sequestration on Wisconsin Conservation Reserve Program land. J. Soil Water Conserv. 58: 58–62.

Izaurralde, R.C., Nyborg, M., Solberg, E.D., Janzen, H.H., Arshad, M.A., Malhi, S.S., Molina-Ayala, M., 1997. Carbon storage in eroded soils after five years of reclamation techniques. In: Lal, R., Kimble, J.M., Follett, R.F., Stewart, B.A. (Eds.), Advances in Soil Science. Soil Processes and the Carbon Cycle. Lewis Publishers, CRC Press, Boca Raton, FL, pp. 369±385.

Lal, R., 2004. Soil carbon sequestration to mitigate climate change. Geoderma, 123(1-2), 1-22.

Lobsey, C. R., Viscarra Rossel, R. A., Roudier, P., Hedley, C. B., 2017. rs-local data-mines information from spectral libraries to improve local calibrations. European Journal of Soil Science,68(6), 840-852.

Minasny, B., Malone, B. P., Mcbratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., ... & Field, D. J., 2017. Soil carbon 4 per mille. Geoderma, 292, 59-86.

Morgan C.L.S., Waiser T.H., Brown D.J., Hallmark C.T., 2009. Simulated in situ characterization of soil organic and inorganic carbon using visible near infrared diffuse reflectance spectroscopy. Geoderma, 151, pp. 249-256

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., ... & Dardenne, P., 2015. Soil spectroscopy: An alternative to wet chemistry for soil monitoring. In Advances in agronomy (Vol. 132, pp. 139-159). Academic Press.

O'Rourke, S.M., Holden, N.M., 2011. Optical sensing and chemometric analysis of soil organic carbon - a cost effective alternative to conventional laboratory methods? Soil Use Manage. 27, 143e155.

Paustian, K., Collins, H.P., Paul, E.A., 1997. Management controls on soil carbon. In: Paul, E.A., Paustian, K., Elliott, E.T., Cole, C.V. (Eds.), Soil Organic Matter in Temperate Agroecosystems. CRC Press, Boca Raton, FL, USA, pp. 15–49.

Reeves I., J., McCarty, G., Mimmo, T., 2002. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. Environmental pollution, 116, S277-S284.

Sørensen L.K and Dalsgaard S., 2005. Determination of clay and other soil properties by near infrared spectroscopy Soil Science Society of America Journal,69, pp. 159-167

Shi, T., Chen, Y., Liu, Y., Wu, G., 2014. Visible and near-infrared reflectance spectroscopy—An alternative for monitoring soil contamination by heavy metals. Journal of hazardous materials, 265, 166-176

Sheppard, N., Willis, H.A. Rigg, J.C., 1985. International Union of Pure and Applied Chemistry (IUPAC). Names, symbols, definitions, and units of quantities in optical spectroscopy (Recommendations 1984). Pure & Applied Chemistry, 57, 105–120.

Staley, T.E., Edwards, W.M., Scott, C.L., Owens, L.B., 1988. Soil microbial biomass and organic component alterations in a notillage chronosequence. Soil Sci. Soc. Am. J. 52, 998–1005.

Stevens, A., van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B., Ben-Dor, E., 2008. Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. Geoderma, 144(1-2), 395-404.

Smith, P., 2004. How long before a change in soil organic carbon can be detected?, Global Change Biology, 10, 1878–1883.

Veronique B.-M., Fernandez-Ahumada E., Palagos B., Roger J.M., McBratney, A.B., 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy TrAc-Trends Anal. Chem., 29, pp. 1073-1081. View Record in Scopus

Virto, I., Barré, P., Burlot, A., Chenu, C. (2012). Carbon input differences as the main factor explaining the variability in soil organic C storage in no-tilled compared to inversion tilled agrosystems. Biogeochemistry, 108(1-3), 17-26.

Viscarra Rossel R.A., Adamchuk, V. I., Sudduth, K. A., McKenzie, N. J., Lobsey, C., 2011. Proximal soil sensing: an effective approach for soil measurements in space and time. In Advances in agronomy (Vol. 113, pp. 243-291). Academic Press.

Viscarra Rossel R.A., Brus, D. J., Lobsey, C., Shi, Z., McLachlan, G., 2016. Baseline estimates of soil organic carbon by proximal sensing: comparing design-based, model-assisted and model-based inference. Geoderma, 265, 152-163.

Viscarra Rossel, R.A., Walvoort, D.J.J., 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.

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

West, T. O. Post, W. M., 2002. Soil organic carbon sequestration rates by tillage and crop rotation: a global data analysis. Soil Sci. Soc. Am. J. 66: 1 930–1 946.

Xie H.T., Yang X.M., Drury C.F. Yang, J.Y., Zhang X.D., 2011. Predicting soil organic carbon and total nitrogen using mid- and near-infrared spectra for Brookston clay loam soil in Southwestern Ontario, Canada. Canadian Journal of Soil Science, 91, pp. 53-63.

CAP.4

CHECKING THE EFFECTS OF THINNING TREATMENTS ON LITTER IN A DEGRADED MEDITERRANEAN PINE FOREST BY VISNIR

4.1.Introduction

Since late 19th century pines were widely used for land restoration in Mediterranean area, as stress-tolerant and pioneer species, but now 31% of Italian Mediterranean pine forests suffers degradation.

In a degraded forest the emissions due to the decomposition of remaining plant material are no more balanced by C stored in woody biomass and soil. So, degradation affects the forest ability in C sequestration and studies are needed to investigate how innovative managements could allow to recover forest functionalities.

In particular, in the forest ecosystem the litter plays a very important role in C fluxes and stock. The rationale of this study deals with the idea that thinning, affecting the amount and composition of litter, may alter the decomposition process, with effects on CO2 emissions, nutrient and C stock in the soil.

Litter is one of the five carbon (C) pools identified by the Intergovernmental Panel on Climate Change (IPCC) as relevant for estimating carbon stock change in terrestrial ecosystems (IPCC Good Practice Guidance for LULUCF). Globally, forest management practices have recognized to have an important role on litter C stock together with the vegetation type and site conditions (Berg and McClaugherty 2008, IPCC, 2006). Nevertheless, the importance of litter for C budget relies on its decomposition, which i) is one of the major processes influencing C fluxes between the terrestrial biosphere and the atmosphere, though carbon dioxide (CO2) emission and ii) it contributes to the formation and stabilization of soil organic matter (SOM), releasing organic compounds into the soil (Oades, 1988; Liski et al. 2002). Moreover, litter dynamics is the main pathways for nutrients input to the soil (Maguire, 1994), contributing up to more than 70% of the annual N input via litter fall (Bauer et al., 2000).

Litter decomposition is an enzyme-mediated biological process carried out by bacteria and fungi. During the progress of litter degradation, a differentiation of enzyme activities and functional diversity have been found, following the accumulation of recalcitrant compounds (Alarcón-Gutiérrez, 2009). An analysis of hydrolytic enzyme potential activities related to the cycling of C, N, P and S is expected to give an insight if different decomposition processes are occurring.

Some authors found good correlations between enzymatic activity and soil spectral properties (Zornoza et al. 2008; Dick et al. 2013, Rinnan and Rinnan, 2007). A spectrum on VisNIR region stores information on organic and inorganic materials in the soil as the radiation will cause individual molecular bonds to vibrate, either by bending or stretching, and they will adsorb light, to various degrees, with a specific energy quantum. In particular, minerals that contain iron (e.g., haematite, goethite) and soil organic matter produces absorptions in the visible

region (400–780 nm) (Mortimore et al., 2004; Sherman and Waite, 1985). The strongest absorbers in the NIR region are the bonds O–H, such as in water, and bonds such as C–N, N–H and CO, characteristic of organic matter. For this reason, spectral properties can be used to identify whether some differences in the presence and amount of specific chemical bonds exist among compared samples.

The study is conducted in a degraded pine forest after thinning treatments were applied to recover ecosystem productivity and forest regeneration. Two different thinning treatments have been compared to un-managed forest, thus obtaining a picture of the best management options able to mitigate climate change through increasing C sequestration and containing green-house gas (GHG) emissions. The objective of the study was a comparison between the spectral properties in VisNIR domain and a set of biochemical properties, over their ability to highlight if some change occurred in litters when different thinnings were applied. Our results aim to embody a contribution to understanding the complex relationships between physico-chemical and biological factors affecting organic matter transformations of especially litter fractions. It belongs to a wider study carried out in the wider frame of the FoResMit LIFe project.

4.2 Materials and methods

4.2.1 Experimental site

The study site is located within the peri-urban forest of Monte Morello (43°51′N; 11°14′E) in the Sesto Fiorentino municipality and close to the urban area of Florence (Italy), at an altitude of about 600 m a.s.l. (fig.4.1).

The climate is typically Mediterranean, with a dry summer in which July is the driest month, while October and November are the rainiest months. During the last decades (from the early 1980s) the total annual rainfall was 1003 mm, concentrated especially in the period from autumn to early spring, and the average annual temperature was 13.9 °C. The basic stone is a calcareous flysch (turbidites) constituted by alternating limestones, marly limestones (“alberese”) marls, claystones and, subordinately, sandstones. Soil is mainly calcareous, with pH ranging between 7.0 and 8.2.

This forest is the result of the reforestations activities realized from 1909 to 1980; specifically, experimental plots are 50 to 60 years old. The main tree species used are Pinus nigra J.F. Arnold, Pinus brutia Ten. subsp. brutia, Cupressus sempervirens, Fraxinus ornus L., Quercus cerris L. and Quercus pubescens L.. After reforestation, the stands have been abandoned with negative consequences on trees stability, high mortality, absence of regeneration, marked susceptibility to adversities and increase of fire risk (Cenni et al., 1998).

Between September and December 2016, three silvicultural treatments have been applied in nine demonstrative plots of 1.5 ha approx (three replicates for each silvicultural option), with the objective of restoring the ecological stability and enhance the resistance and resilience of the forest (Paletto et al., 2017). Within each plot two monitoring sub-plots (circular fixed-area of 531 m^2) has been selected. Thinning treatments were based on three silvicultural options: traditional thinning, selective thinning and absence of treatment (control). The traditional silvicultural

treatment was based on a low thinning, otherwise known as ‘thinning from below’ and was of medium-heavy intensity, removing most of the dominated trees and including also some trees of the dominant layer. With the selective thinning, the best trees of the stand were selected according to vigour and stability (“positive selection”) and their growth and development was actively promoted by removing competitors in the dominant layer, whereas plants in the dominated layer ware harvested only in case of economic convenience. Moreover, considering that pines are very light demanding species, all the suppressed and sub-dominant trees were removed with the aim to avoid the increasing deadwood with time (fig. 4.2).

Figure 4.1: images from Montemorello forest (on the left) and its litter and soil profile (on the right) (by Lorenzetti et al., 2018).

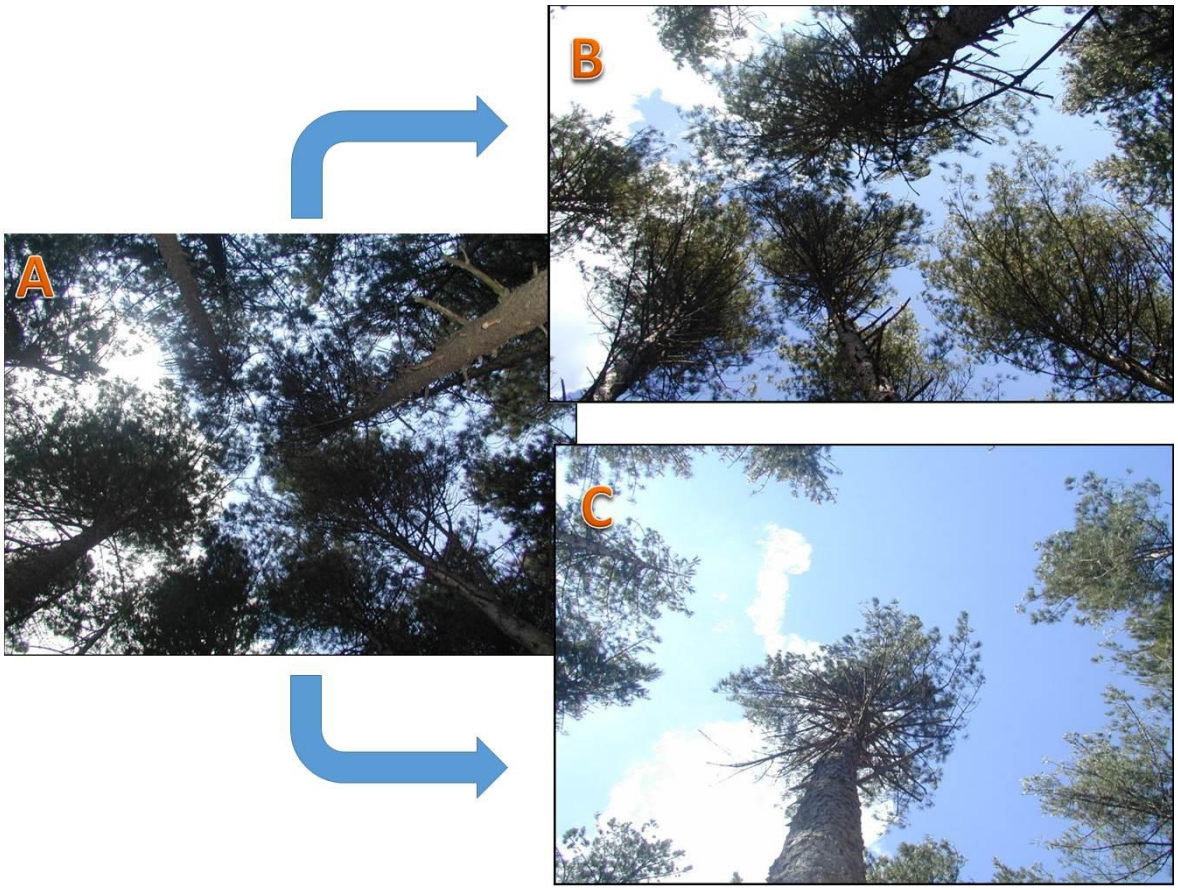


Figure 4.2. a) Control (no thinning); b) Traditional thinning (thinning from below), c) Innovative selective thinning (growth and development of best trees are actively promoted by removing all directs competitors) (by http://lifeforesmit.com).

4.2.2 Sampling and characterization

Forest floor has been collected in January 2017 in each sub-plot, one month after the end of harvesting, for a total of 18 sampling points, by pressing a 600-700 cm² steel sheet sampling frame 10 cm deep (or similar) into the forest floor and collecting all litter material above the mineral soil, following the approach of Kavvadias et al. (2001).

Samples from the two sub-plots have been merged together and the plot was the unit of replication.

The samples were transported to the laboratory, dried and then fractionated by hand sorting and sieving. In each of these samples the horizons L, F and H were separated into the following three fractions. The L horizon is composed of fresh or slightly discoloured, with no or weak breaking up, material. It contains freshly fallen leaves, needles, twigs, cones, bark chunks, dead moss, dead lichens, dead herbaceous stems, and flower parts in various stages of decay but still recognizable as individual plant parts (i.e., visible fibrous materials). The F horizon is composed of medium to strongly fragmented material with many mycelia and thin roots and the H horizon is a humified amorphous material, containing highly decomposed (i.e.,

unrecognizable) plants parts (Kavvadias et al. 2001; Hoosbeek and Scarascia-Mugnozza 2009). Mineral soil was removed through successive sieving (10 – 5 and 2 mm mesh stainless steel sieves).

4.2.3 Wet chemistry analyses

C and N were determined with a cut-mill for carbon and nitrogen content determination by dry combustion with an elemental analyzer (Flash 2000, Thermo Fisher).

Enzyme activity was measured according to the methods of Marx et al. (2001) and Vepsäläinen et al. (2001), based on the use of fluorogenic methylumbelliferyl (MUF)-substrates. Homogenized forest floor fractions were analysed for N-acetyl-β-glucosaminidase (NAG), β-glucosidase (βG), butyrate esterase (BUT), acid phosphatase (AP), arylsulphatase (ARYL), β-xylosidase (XYL), cellulose (CELL) and acetate esterase (AC) activity using MUF conjugated surrogate substrates (Sigma, St Louis, MO, USA). A homogenous suspension was obtained by homogenising 2 g samples with 50 mL deionized water with UltraTurrax at 9600 rev / min for 3 min. Aliquots of 50 µL were withdrawn and dispensed into a 96 well microplate (3 analytical replicates/sample/substrate). 50 µL of Na-acetate buffer pH 5.5 was added to each well. Finally, 100 µL of 1 mM substrate solution were added giving a final substrate concentration of 500 µM. Fluorescence was measured after 0, 30, 60, 120, 180 min of incubation at 30 °C. Fluorescence (excitation 360 nm; emission 450 nm) was measured with an automated fluorimetric plate-reader (Fluoroskan Ascent).

The order of magnitude of the values obtained for the different enzymatic responses varies considerably depending on the specific activity being determined, thus leading to some enzyme having more weight than others. To resolve this problem, the sum of the percentage of the maximum value found for a specific enzymatic response across all enzymes was used for the calculation of the sum of enzymes (SUM).

4.2.4. Spectral data

Data were collected on the samples homogenized at 0.5 mm. Each spectrum was made up of 1,151 wavelengths, from 2,500 to 200 nm. Vis–NIR spectra were recorded as percent reflectance (R%). Data acquisition was carried out by means of the Jasco software (Model VWTS-581 version 2.00A) and a spectralon surface was use as white reference.

4.2.5. Statistical analysis

Discriminant function analysis (DA) was performed for the litter fractions. Squared Mahalanobis distances between group centroids were determined. Two significant discriminatory roots were derived, and the results of DA were graphically presented in two dimensions.

Soil enzymatic activities were used as biochemical grouping variables.

DA was also performed on the spectral data as grouping variables. In order to reduce the variable size a principal component analysis (PCA) was carried out, 30 principal components (PCs) were extracted and selected to be used as grouping variables. Only the PCs with the higher

discriminating power were used. They were selected based on the Partial lambda. DA was performed over each fraction separately, and over all the samples together.

PCA was performed by Unscrumble software.

4.3. Results

4.3.1 Forest floor composition and characterization

Forest floor biomass showed the maximum amount in F fraction and the minimum in L (Table 4.1). This composition was affected by treatments, which increased significantly with traditional thinning in F fraction. The three components of forest floor showed well distinguished characteristics, with a decrease of C content along the litter decomposition levels, with a significantly lower C % in H fraction with respect to L and F (Table 1). A correspondent increase of N % determined a decrease of C/N ratio, both significantly different in H fraction. This pattern occurred independent of treatments, which didn't affect significantly any fraction.

Table 4.1. mean values of Biomass, C and N percentage and C/N ratio of forest floor.

		Biomass (kg m^{-2})	C (%)	N (%)	C/N ratio
L	Control	0.32 ^{e}	42.1 ^{a}	0.7 ^{b}	64.0 ^{a}
	Selective	0.43 ^{e}	40.5 ^{a}	0.6 ^{b}	64.3 ^{a}
	Traditional	0.68 ^{de}	42.3 ^{a}	0.8 ^{b}	59.6 ^{a}
F	Control	1.42 ^{bc}	38.5 ^{a}	0.8 ^{b}	47.7 ^{a}
	Selective	1.93 ^{ab}	38.0 ^{a}	0.7 ^{b}	57.3 ^{a}
	Traditional	2.19 ^{a}	38.8 ^{a}	0.7 ^{b}	55.8 ^{a}
H	Control	0.96 ^{cde}	26.3 ^{b}	1.1 ^{a}	23.2 ^{b}
	Selective	1.17 ^{cd}	23.7 ^{b}	1.1 ^{a}	21.3 ^{b}
	Traditional	1.34 ^{bcd}	24.7 ^{b}	1.1 ^{a}	22.2 ^{b}
ANOVA	Litter fraction	0.000	0.000	0.000	0.000
	Treatment	0.053	n.s.	n.s.	n.s.
	Treatment*fraction	n.s.	n.s.	n.s.	n.s.

4.3.2. Enzyme activities

Overall, enzyme activities changed significantly in the three litter fractions (Table 4.2), which showed a significantly different pattern (Fig. 4.3). H fraction showed the highest activities of the two esterases (AC and BUT) and, with a six fold increase, ARYL. L and F fraction showed the highest activities of AP, cellulose (CELL and βG), hemicellulose (XYL) and chitin (NAG) degrading enzymes.

Selective thinning reduced all enzyme activities with respect to Control and Traditional treatment in the three fractions, significantly for cellulose and hemicellulose degrading enzymes (CELL, βG, XYL) in L fraction between the two treatments (Table 4.2).

Table 4.2. Enzyme activities in the three litter fractions for the three thinning treatments, expressed as nmol mub g-1 h-1.

	CELL	AP	β-G	NAG	XYL	BUT	AC	ARYL	SUM
L									
Control	175 ^{ab}	1477 ^{a}	646 ^{abc}	248 ^{ab}	97 ^{abc}	699 ^{bd}	815 ^{ab}	11 ^{b}	364 ^{ab}
Selective	115^{b}	1069 ^{ab}	502 ^{c}	203 ^{b}	56 ^{c}	477 ^{d}	668 ^{b}	9 ^{b}	263^{b}
Traditional	247^{a}	1411 ^{a}	854 ^{a}	289 ^{ab}	112^{a}	707 ^{bcd}	881 ^{ab}	13 ^{b}	415^{a}
F									
Control	230^{ab}	1385 ^{a}	812 ^{ab}	318 ^{ab}	105 ^{ab}	1013 ^{abc}	1052 ^{ab}	32 ^{b}	434^{a}
Selective	168 ^{ab}	1225 ^{ab}	631 ^{abc}	267 ^{ab}	83 ^{abc}	831 ^{bcd}	1064 ^{ab}	22 ^{b}	404 ^{ab}
Traditional	223 ^{ab}	1326 ^{a}	811 ^{ab}	385 ^{a}	100 ^{abc}	849 ^{bcd}	1056 ^{ab}	31 ^{b}	461^{a}
H									
Control	204 ^{ab}	1077 ^{ab}	714 ^{abc}	272 ^{ab}	88 ^{abc}	1407 ^{a}	1274 ^{a}	163 ^{a}	485^{a}
Selective	118^{b}	830 ^{b}	506 ^{c}	190 ^{b}	64 ^{bc}	1016 ^{abc}	1061 ^{ab}	152 ^{a}	367 ^{ab}
Traditional	158 ^{ab}	1047 ^{ab}	566^{bc}	220 ^{b}	74 ^{abc}	1115 ^{ab}	1122 ^{ab}	169 ^{a}	421 ^{a}
					Analysis of variance				
Litter fraction	n.s.	0.028	0.168	0.05	0.270	0,001	0.084	0.000	0.012
Treatment	0.07	0.109	0.041	0.13	0.05	0.09	n.s.	n.s.	n.s.
Treatment*Litter fraction	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

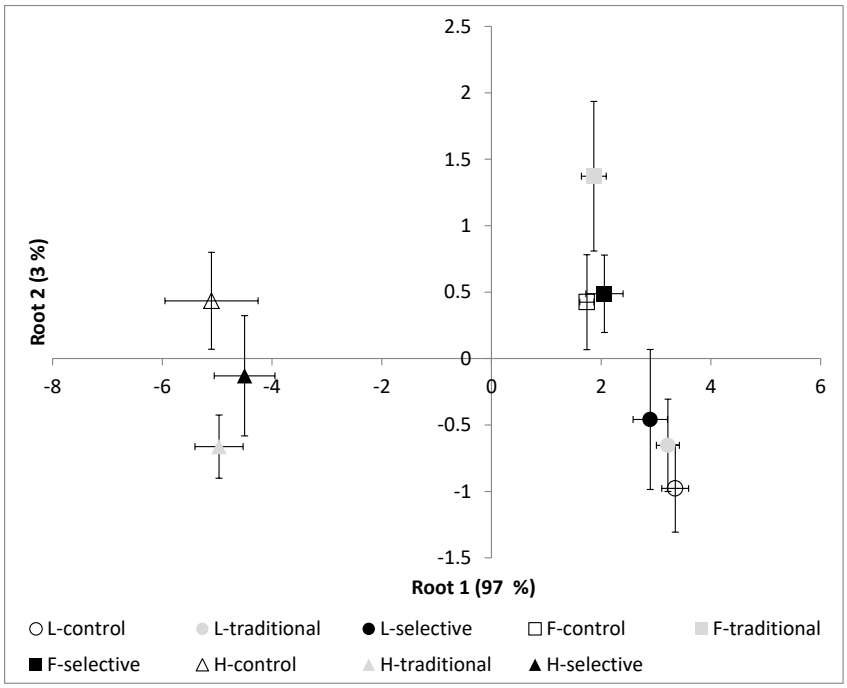


Figure 4.3. Discriminant Analysis of enzyme activities, showing separation among litter fractions and treatments within each fraction.

4.3.3. Discriminating thinning effect by VisNIR reflectance

The 30 PCs used for the DA explained as a whole 99.97% of the spectral variance, 88.7% of which was explained by the first one.

By performing DA over all samples a sub set of 9 PCs was selected as those with the highest discriminating power. Among them, the subset composed of the PCs number 6, 9 and 25 gave the best results when all the data were used; the PCs 6, 22, and 25 gave the best results over the L fraction, 2, 22, 26 over the H fraction and 6, 9, 27 over the F.

As reported in the table 4.3, the DA highlighted significant differences in the spectral properties of the forest floor under the three thinning conditions. That was true when the DA was performed without taking into account differences among the fractions, but also for the singles fraction, except for the F, where the spectral properties demonstrated the lowest discriminating power over the thinning groups. In all cases control and traditional thinning seemed to be more similar. Conversely, selective thinning was the most distant population. Results were graphically summarized in figure 4.4.

PC6 was correlated with enzyme activities (CELL, bG, XYL, AP) in the whole forest floor and further with BUT, ARYL, C and N percentage in L fraction (Table 4.4). ARYL was correlated also with PC22 of L fraction. Regarding F fraction, XYL and ARYL were positively correlated with PC6 and AP and XYL negatively with PC27. PCs of H fraction were not correlated with any of the analyzed parameters.

Table 4.3. Results from discriminant Function Analyses of VisNIR reflectance for the forest floor as a whole (LFH), and for the single fractions (p-level *<0.05; **<0.001; ***<0.0001; °<0.1).

Fractions	Forest floor	L	F	H
Variables in the model (p-level refer to F-remove value):	PC6** PC 9* PC 25*	PC 6** PC 22 PC 25°	PC 6 PC 9* PC 27	PC 2* PC 22* PC 26*
n. of groups:	3	3	3	3
n. samples	54	18	18	18
Wilks' Lambda	0.571***	0.265*	0.424 °	0.241*
% of correct Predicted classifications				
Control	56	67	33	83
Traditional thinning	50	50	50	100
Selective thinning	89	100	83	83
Tot	65	72	56	89
Squared Mahalanobis Distance				
Control – Traditional t.	0.12	1.00	0.34	4.11*
Control - Selective t.	3.10***	9.17*	4.51*	5.97*
Traditional t. - Selective t.	3.02***	7.62*	4.71*	5.69*

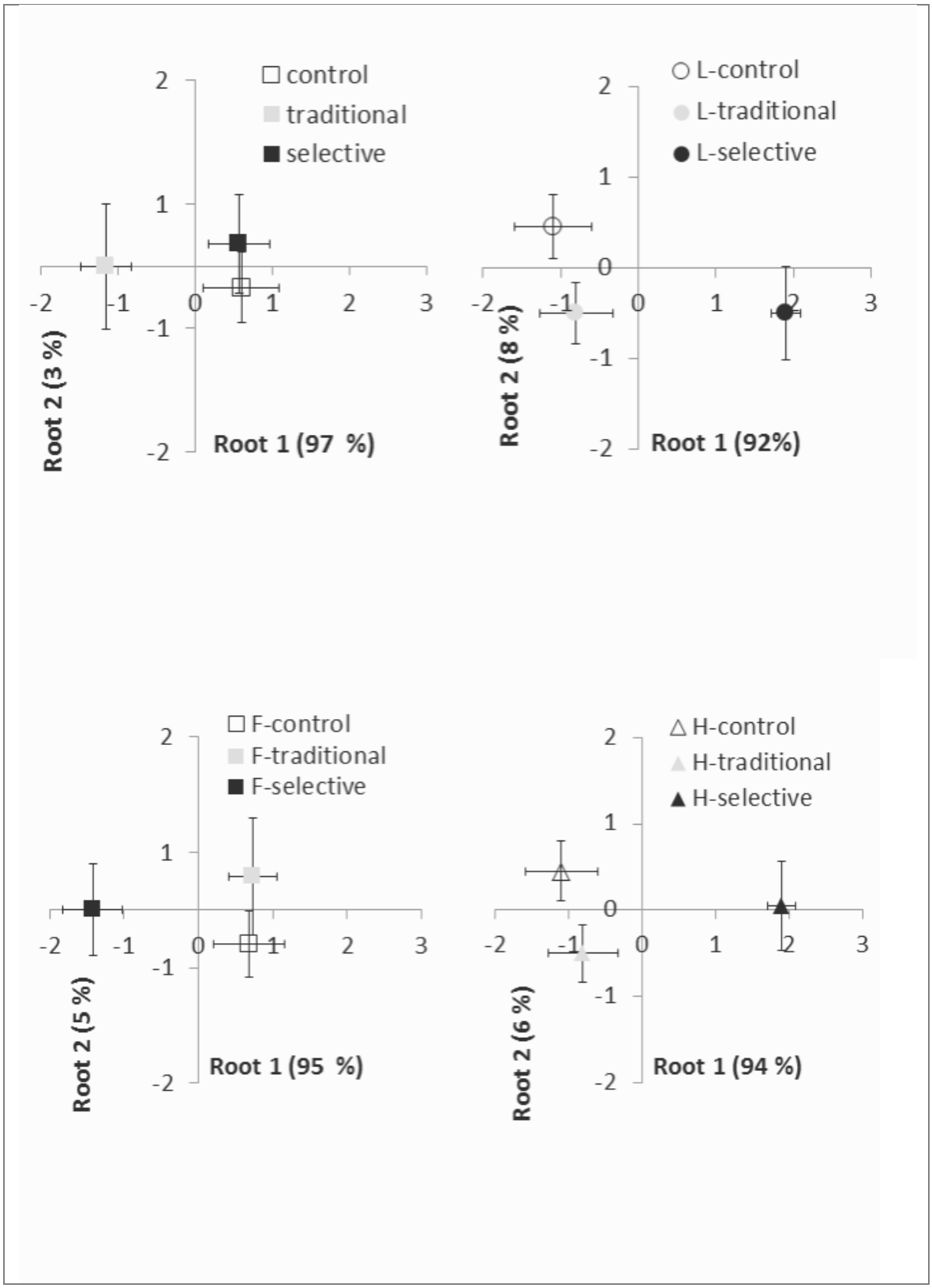

control
traditional
selective
Root 2 (3 %)
Root 1 (97 %)
L-control
L-traditional
L-selective
Root 2 (8 %)
Root 1 (92%)
F-control
F-traditional
F-selective
Root 2 (5 %)
Root 1 (95 %)
H-control
H-traditional
H-selective
Root 2 (6 %)
Root 1 (94 %)

Figure 4.4: DAs of spectral properties, showing separation among treatments over the all forest floor.

Table 4.4: correlation between biochemical properties and spectral properties within the fraction and over all the forest floor samples.

		CELL	AP	βG	XYL	BUT	ARYL	N (%)	C (%)
Forest Floor	PC6	0.527**	0.598**	0.566**	0.595**	n.s.	n.s.	n.s.	n.s.
L	PC6	0.685*	0.679*	0.688*	0.655*	0.6936*	0.620*	0.705*	0.754*
	PC22	n.s.	n.s.	n.s.	n.s.	n.s.	0.773*	n.s.	n.s.
F	PC6	n.s.	n.s.	n.s.	0.660*	n.s.	0.692*	n.s.	n.s.
	PC27	n.s.	-.641*	n.s.	-0.673*	n.s.	n.s.	n.s.	n.s.

4.4. Discussion

4.4.1 Forest floor composition and characterization

The range of litter amount in the degraded pine forest were in line with those observed for other coniferous forests in temperate and Mediterranean environments (Florence and Lamb, 1973; Brovkin et al., 2011). Thus, forest degradation seemed to affect more deadwood than litter amount (De Meo et al., 2017). Also, the distribution of litter was similar to that observed by other studies (e.g. Florence and Lamb, 1973; Rodkey et al., 1994), who reported F-layer constitutes the greater part of the total litter present in most stands. In our case, F fraction accounted for 53 % of total forest floor biomass, which included the most differentiated litter pool.

Overall, forest floor fractions were well characterized in terms of chemical composition, with an evident decrease of C and increase of N as decomposition advance. This pattern is well-known either in coniferous and broadleaves species (see the review of Berg, 2014). C is lost first during decomposition process, corresponding well with a parallel decrease of cellulose and hemicellulose degrading enzymes (Sňajdr et al., 2011). Our data did not show a clear trend of CELL, □G and XYL activities, which were similar in the three fractions. Indeed the decrease of C was more related to a decrease of AP and an increase of esterases and ARYL activities during decomposition. AC and BUT are unspecific esterases, which are involved in the cycling of carbon (Tabatabai and Fu 1992). BUT activity is oriented toward the more recalcitrant compounds, increasing in samples rich in recalcitrant compounds of short alkyl chains, generated during the course of litter degradation (Boczar et al., 2001; Allison & Vitousek, 2005). Thus, an increase in H fraction suggested significant changes in organic matter chemical composition, as a consequence of selective degradation by soil microorganisms, with the accumulation of more

recalcitrant alkyl C groups, as observed by more specific NMR analyses (Baldock et al., 1992; Quideau et al., 2000; Kavdir et al., 2005).

4.4.2 Thinning effects

Thinning operations basically increased litter biomass in all fractions, and the highest increase was observed in F fraction with traditional thinning (thinning from below). Either trees cutting or removal might have favored accumulation and physical break up of litter, since a chemical evolution from L to H fraction within 1 month is not likely. Also, C and N content of forest floor fractions was not affected by treatments, confirming similar characteristics in the three thesis.

The increase of biomass did not correspond to an increase of forest floor activity. Enzyme activities showed the lowest values in the selective thinning in all fraction, showing a clear trend towards lower decomposition, mainly for cellulose and hemicellulose degrading enzymes. Lower enzyme and decomposition rates were observed after thinning in different ecosystems (Waldrop et al., 2003; Lindo & Visser, 2003; Wang et al., 2013; Akburak and Makineci, 2016). These results supported a possible deceleration of litter decomposition following thinning operations in the short term.

Results from the enzymes analysis were upheld by the spectral response. The bigger distance of the samples under selective thinning from other groups, confirmed the presence of an effect of the selective thinning on the litter, especially in the most superficial fraction. Actually, in the L fraction, in particular, the discriminant variables showed a correlation with the enzymatic properties. While, this correlation was reduced only to one enzyme in the F fraction.

We hypothesized that decreased litter decomposition after thinning might be related more to a physical effect than to biochemical changes. In fact, litter composition did not change with thinning, whereas the increase in solar radiation and soil compaction might have affected litter decomposition (Rey et al., 2002; Tan et al., 2008).

However, the spectroscopic analysis showed a change due to the different thinning, much more than the chemical properties and enzymatic activity. Actually, VisNIR gave integrated information on physical, chemical, and biological feature of the samples. It can justify the high power in detaching slight differences due to effect of a very recent changing in the forest management.

Even in the H fraction, where no evidence was pointed out by the other analysis the spectral information highlighted a separation among treatments. Differently from the L fraction, for H the discriminant spectral variables did not have any correlation with the analytical data. It could mean that the spectral difference in H should be attributed to other not investigated properties of the samples. We may suppose, for instance, that some change occurred in this deeper fraction after the treatment as a consequence of the death of the superficial roots that generally occurred immediately after the thinning.

4.5. Conclusions

Overall, decomposition process was characterized by changes in biochemical among litter fractions.

The effect of thinning of soil reduced activity of litter decomposition could be maily consider a mechanical consequence. Further studies on the role of physical compaction on forest floor might increase our knowledge of thinning effects on litter composition and activity, therefore giving more precise indications on management strategies.

Although findings on enzymatic activity and chemical parameters didn't support the presence of significant differences community composition among the three forest management type, it was not true for VisNIR spectral data. Since, it gave integrated information on physical, chemical, and biological feature of the sample, it demostrated a higher power in detaching slight differences due to effect of a very recent changing in the forest management. L fraction showed clearer correlation with the type of thinning than F, where no significative spectral difference were highlighted among the treatments. As no correlation was found between main feature in VisNIR domain and chemical and biochemical properties, spectral difference in H fraction could be attributed to other not investigated properties of the samples, such as the death of the superficial roots generally occurring immediately after the thinning.

Due to the presence of some spectral differences among the three investigated forest management techniques, further study of longer time of observation may allow to reveal some more positive responses also for the chemical and biochemical properties.

Acknowledge

The work was financially supported by the LIFE program, in the context of FoResMit project (LIFE14/CCM/IT/905) "recovery of degraded coniferous Forests for environmental sustainability Restoration and climate change Mitigation".

Thank to Alessandra Lagomarsino, Alssandro Elio Agnelli, and to the other authors of this study th et allowed me to carry out the spectral aquisitions over litter samples yet collected, prepared and analyzed with traditional chemical and enzimatic analyses.

Reference

Akburak, S., Makineci, E., 2016. Thinning effects on soil and microbial respiration in a coppice-originated Carpinus betulus L. stand in Turkey. iForest 9:783-790. doi: 10.3832/ifor1810-009.

Alarcón-Gutiérrez, E., Floch, C., Augur, C., Le Petit, J., Ziarelli, F., Criquet, C., 2009. Spatial variations of chemical composition, microbial functional diversity, and enzyme activities in a Mediterranean litter (Quercus ilex L.) profile. Pedobiologia 52 (6):387-399.

Allison, S.D., Vitousek, P.M., 2005. Responses of extracellular enzymes to simple and complex nutrient inputs. Soil Biol Biochem 37 (5):937-944.

Baldock, J.A., Oades, J.M., Waters, A.G., Peng, X., Vassallo, A.M., Wilson, M.A., 1992. Aspects of the chemical structure of soil organic materials as revealed by solid-state 13C NMR spectroscopy. Biogeochemistry 16 (1):1–42.

Bauer, G.A., Gebauer, G., Harrison, A.F., Högberg, P., Högbom, L., Schinkel, H., Taylor, A.F.S., Novak, M., Buzek, F., Harkness, D., Person, T., Schulze, E.D., 2000. Biotic and Abiotic Controls Over Ecosystem Cycling of Stable Natural Nitrogen, Carbon and Sulphur Isotopes. In: Schulze ED (ed) Carbon and Nitrogen Cycling in European Forest Ecosystems. Ecological Studies (Analysis and Synthesis), vol 142. Springer, Berlin, Heidelberg.

Berg, B., 2014. Decomposition patterns for foliar litter – A theory for influencing factors. Soil Biol Biochem 78:222-232.

Berg, B., McClaugherty, C. (eds), 2008. Plant Litter. Decomposition, Humus Formation, Carbon Sequestration. Springer-Verlag Berlin Heidelberg.

Boczar, B.A., Forney, L.J., Begley, W.M., Larson, R.J., Federle, T.W., 2001. Characterization and Distribution of Esterase Activity in Activated Sludge. Water Res 35 (17):4208-4216

Brovkin, V., van Bodegom, P.M., Kleinen, T., Wirth, C., Cornwell, W.K., Cornelissen, J.H.C., Kattge, J., 2012. Plant-driven variation in decomposition rates improves projections of global litter stock distribution. Biogeosciences 9:565–576.

Cenni, E., Bussotti, F., Galeotti, L., 1998. The decline of a Pinus nigra Arn. Reforestation stand on a limestone substrate: The role of nutritional factors examined by means of foliar diagnosis. Ann Sci Forest 55:567–576.

De Meo, I., Agnelli, A.E., Graziani, A., Kitikidou, K., Lagomarsino, A., Milios, E., Radoglou, K., Paletto, A., 2017. Deadwood volume assessment in Calabrian pine (Pinus brutia Ten.) peri-urban forests: Comparison between two sampling methods. J Sustain Forest 36(7):666-686.

Dick, W.A., Thavamani, B., Conley, S., Blaisdell, R., Sengupta, A.., 2013. Prediction of β- glucosidase and β-glucosaminidase activities, soil organic C, and amino sugar N in a diverse population of soils using near infrared reflectance spectroscopy. Soil Biol Biochem 56:99-104.

Florence, R.G., Lamb, D., 1973. Influence of stand and site on radiata pine litter in South Australia. New Zealand. J Forest Sci 4:502-510.

Hoosbeek, M.R., Scarascia-Mugnozza, G.E., 2009. Increased litter build up and soil organic matter stabilization in a poplar plantation after 6 years of atmospheric CO2 enrichment (FACE): final results of POP-EuroFACE compared to other forest FACE experiments. Ecosystems 12(2):220-239.

Intergovernmental Panel on Climate Change (IPCC), 2006. IPCC Guidelines for National Green-house Gas Inventories, vol. 4, Agriculture, Forestry and Other Land Use. Eggleston S. (ed) Inst. for Global Environ. Strategies, Hayama, Japan.

Kavdır, Y., Ekinci, H., Yüksel, O., Mermut, A.R., 2005. Soil aggregate stability and 13C CP/MAS-NMR assessment of organic matter in soils influenced by forest wildfires in Canakkale, Turkey. Geoderma 129(3-4):219-229.

Kavvadias, V.A., Alifragis, D., Tsiontsis, A., Brofas, G., Stamatelos, G., 2001. Litterfall, litter accumulation and litter decomposition rates in four forest ecosystems in northern Greece. Forest Ecol Manag 144(1-3):113-127.

Lindo, Z., Visser, S., 2003. Microbial biomass, nitrogen and phosphorus mineralization, and mesofauna in boreal conifer and deciduous forest floors following partial and clear-cut harvesting. Can J Forest Res 33(9):1610-1620.

Liski, J., Perruchoud, D., Karjalainen, T., 2002. Increasing carbon stocks in the forest soils of western Europe. Forest Ecol Manag 169(1-2):159-175.

Lorenzetti, R., Guerrero, C., Di Iorio, E., Agnelli E.A, Colombo, C., Lagomarsino, A., 2018. Discriminating the effects of forest management on litter and soil ina medioterranean pine forest by VisNIR. In: ATTI DELLA ESSC Conference "Soil and Water Security: challenges for the next 30 years!". Imola, 6-8/06/2018.

Maguire, D.A., 1994. Branch mortality and potential litterfall from Douglas-fir trees in stands of varying density. Forest Ecol Manag 70(1-3):41-53.

Marx, M.C., Wood, M., Jarvis, S.C., 2001. A microplate fluorimetric assay for the study of enzyme diversity in soils. Soil Biol Bbiochem 33(12-13):1633-1640

Mortimore, J.L., Marshall, L.J.R., Almond, M.J., Hollins, P., Matthews, W., 2004. Analysis of red and yellow ochre samples from Clearwell Caves and Çatalhöyük by vibrational spectroscopy and other techniques. Spectrochim Acta A 60(5):1179- 1188.

Oades, J.M., 1988. The retention of organic matter in soils. Biogeochemistry 5(1):35-70.

Paletto, A., De Meo, I., Grilli, G., Nikodinoska, N., 2017. Effects of different thinning systems on the economic value of ecosystem services: A case-study in a black pine peri-urban forest in Central Italy. Ann For Res 60.

Quideau, S.A., Anderson, M.A., Graham, R.C., Chadwick, O.A., Trumbore, S.E., 2000. Soil organic matter processes: characterization by 13C NMR and 14C measurements. Forest Ecol Manag 138(1-3):19-27.

Rey, A., Pegoraro, E., Tedeschi, V., De Parri, I., Jarvis, P.G., Valentini, R., 2002. Annual variation in soil respiration and its components in a coppice oak forest in Central Italy. Global Change Biol 8:851-866.

Rinnan, R., Rinnan, Å., 2007. Application of near infrared reflectance (NIR) and fluorescence spectroscopy to analysis of microbiological and chemical properties of arctic soil. Soil Biol Biochem 39:1664-1673.

Rodkey, K.S., Kaczmarek, D.J., Pope, P.E., 1994. The distribution of nitrogen and phosphorus in forest floor layers of oak-hickory forests of varying productivity. In Proceedings 10th Central Hardwood Forest Conference)

Šnajdr, J., Valášková, V., Merhautová, V., Herinková, J., Cajthaml, T., Baldrian, P., 2008. Spatial variability of enzyme activities and microbial biomass in the upper layers of Quercus petraea forest soil. Soil Biol Biochem 40(9):2068-2075.

Sherman, D.M., Waite, T.D., 1985. Electronic spectra of Fe3+ oxides and oxide hydroxides in the near-IR to near-UV. Amer Mineral 70:1262-1269.

Tabatabai, M.A., Fu, M., 1992. Extraction of enzymes from soils. Soil Biochem 7:197-227.

Tan, X., Chang, S.X., Comeau, P.G., Wang, Y., 2008. Thinning effects on microbial biomass, N mineralization, and tree growth in a mid-rotation fire-origin lodgepole pine stand in the lower foothills of Alberta, Canada. Forest Sci 54(4):465-474.

Vepsäläinen, M., Kukkonen, S., Vestberg, M., Sirviö, H., Niemi, R.M., 2001. Application of soil enzyme activity test kit in a field experiment. Soil Biol Biochem 33(12- 13):1665-1672.

Waldrop, M.P., McColla, J.G., Powers, R.F., 2003. Effects of Forest Postharvest Management Practices on Enzyme Activities in Decomposing Litter. Soil Sci Soc Am J 67:1250–1256.

Wang, H., Liu, W., Wang, W., Zu, Y., 2013. Influence of Long-Term Thinning on the Biomass Carbon and Soil Respiration in a Larch (Larix gmelinii) Forest in Northeastern China. The Scientific World Journal.

Zornoza, R., Guerrero, C., Mataix-Solera, J., 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 Biol Biochem 40:1923e1930.

CAP.5

BUILDING THE ITALIAN NATIONAL SOIL SPECTRAL LIBRARY

5.1. Introduction

Successful soil property predictions based on diffuse reflectance spectroscopy (DRS) in the Visible and near infrared (VisNIR) domain were found by several authors (e.g. Ben-Dor and Banin, 1995; Stenberg et al., 1995; Reeves and McCarty, 2001; Shepherd and Walsh, 2002; Dematthê et al., 2004; Brown et al., 2006; Viscarra Rossel et al., 2006). A large number of studies were carried on in several area of the world at local and regional scale since the 80s (e.g. Dalal and Henry, 1986). However, the amount of local experiments deeply increased from the 2000s (Viscarra Rossel et al., 2016). More recently, findings at wider scale were also obtained. Estimation of soil properties at national level started to increase from the 2010s (e.g. Viscarra Rossel and Lark, 2009, Viscarra Rossel and Behrens, 2010, and Viscarra Rossel and Webster, 2012 for Australia; Ji et al. 2015, for China; Goge et al., 2012 for France). However only few countries have currently published their own national soil spectral library: Among them: France (Goge et al., 2012), Danimarca (Knadel et al., 2012), Florida (Vasques et al. 2010), Czech Republic (Brodsky et al., 2011).

Nocita et al. (2015a) provided a wide and exhaustive review of the state of the art of large spectral libraries. Since 2008 a group of scientists from eight countries started to create a Global Soil Spectral Library where 23,631 soil spectra have been currently collected from 92 countries in seven continents (Africa, Antarctica, Asia, Europe with 3518 of which 209 from Italy, North and Central America, Oceania, South America) (Viscarra Rossel et al., 2016).

For European area, a spectral library was available for free: LUCAS (Land Use/Cover Area frame Statistical Survey), consists of about 20,000 topsoil (0 -20 cm) samples, that were collected in order to assess the state of the soil across Europe, under the supervision of the Joint Research Centre of the European Commission (JRC). Its metadata contains 13 soil properties (Stevens et al., 2013).

Since a soil spectral library is an essential step for using Vis-NIR for soil analysis, some studies focused on library creation and management (e.g. Shepherd and Walsh, 2002, Brown et al., 2006, Cellion et al 2009, Bellinaso et al., 2010, Genot et al. 2011, Knadel et al., 2012), and Nocita et al. (2015a) briefly summarized the guide lines for properly building a large soil spectral library. According to these last, soil spectral libraries have to meet the following main requirements: the samples should be representative for the territory in which they should be used (Knadel et al., 2012); library has to be completed with a metadata database that stores the corresponding auxiliary information on the soils: type of material (soil, parent material), sample preparation, location of the sample with geographic coordinates and environmental information, soil classification, soil laboratory measurements – chemical, physical, and potential biological ones, relationally linked together (Brodský, et al. 2011); the reference data used for calibration should be obtained with reliable analytical methods (Knadel et al., 2012); the samples should be

carefully handled and scanned (Knadel et al., 2012); future access to the soil samples for new scanning has to be allowed (Viscarra Rossel et al., 2016).

Once built, a national soil spectral library can fulfill the following functions:

- furthering the research on soil vis–NIR spectroscopy (Viscarra Rossel et al., 2016);
- estimating soil attributes, and functions where measurements are lacking and conventional methods too expensive (Viscarra Rossel et al., 2006; Nocita et al., 2015, Sanchez et al., 2009);
- monitoring soil at scales ranging from regional to global (Viscarra Rossel et al., 2016);
- modelling developments of digital soil mapping applications (Brodský, et al. 2011).

This study aims to provide the first version of the Italian Soil Spectral Library, and the first examples in the calibration of national models. The overall objective is to build a representative soil spectral library of the Italian soils for statistical inference models to allow the exploitation of rapid soil quantitative predictions and classifications for digital soil mapping and monitoring the activities at national or local scale, trying to cover most of the Italian soil variability. Since a large amount of free data are nowadays available, we carried out a comparison between data from Italian library and the free ones for the Italian territory, questioning if the use of a national database could add some useful information for soil investigation.

5.2. Materials and methods

5.2.1 Soil sample

The development of a National soil spectral library for Italy is going to be implemented with the collaboration of the Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria, (CREA Research Center for Agriculture and Environment).

Along the last decades the center handled the building and management of the national soil database (SISI –Soil Information System of Italy), which collects pedological and environmental information about around fifty-six thousand georeferenced and analyzed soil observations in the Italian territory (Costanini et al., 2013). A part of these soil samples is physically stored in the soil archive of the center (pedoteca). About 16,800 analyzed soil samples are currently stored in the soil archive. They are from different studies carried out by the Center during the last decades. The materials used in the library up to this moment, come from these various studies carried out by CREA. Over all the available data a control of quality was carried out, avoiding to export any eventual mistake from ISIS to the spectral library, and providing homogeneity of the metadata information.

The aim of this library is to involve the wide range of soil variations in Italy. Therefore, a huge amount of samples available in the pedoteca were scanned for collecting spectra. In this first building stage, the samples were selected when at least some most common pedological data occurred. In particular, the samples chosen for entering in the library shall have at least the following metadata:

-geographic coordinates (WGS system)

- upper and lower limits
- one among: soil organic carbon (%) (TOC) or texture (% clay, silt, sand).

However, for most of samples, much other physical and chemical properties were available, as:

-pH
- total Carbonate content (%)
- active Carbonate (%)
- total Nitrogen content (‰) (TN)
- cation exchange capacity (CEC) (cmol(+)/kg),
- exchangeable cations (cmol(+)/kg),
- exchangeable Sodium Percentage (ESP) (%)
- bulk density (BD) (g/cm3)
- available water capacity (AWC) (mm/m)
- electrical conductibility (EC) (dS/m)
- stoniness (%)

5.2.2 Metadata

The metadata referring to wet chemistry analyses were subjected to standard laboratory procedure. The soil samples were air-dried, ground and sieved to 2-mm. The basic chemical and physical soil properties were obtained using standard laboratory procedures. The soil pH was measured using a 1:5 (w/v) (ISO 10390 1994). CEC was measured using the method proposed by Bower and Hatcher (Bower & Hatcher 1966 in Klute 1996) The soil carbon content was measured using the Walkley-Black, (1934). $CaCO_3$ content was measured using the volumetric calcimeter method described by Looppert and Suarez (1996). The particle size distribution (fractions of clay, silt, and sand) was obtained by the hydrometer method (Gee & Or 2002).

Spectral signatures inherited all the related geographic data of the samples stored in ISIS. They deal with:

- lithological and morphological information;
- Soil regions and soil systems (Costantini et al., 2013).

5.2.3 Spectra collection

The spectra were acquired by means of FieldSpec 3Hi-Res (ASDi), a portable spectroradiometers manufactured by Analytical Spectral Devices Inc. (ASD, Boulder, CO, USA). FieldSpec instrument was yet adopted by several scientists for applications in areas such as agricultural analysis, field and laboratory mineral and soils analyses. The device works in a spectral range between 350-2500 nm; it is equipped with three detectors: one silicium photodiode array covering 350 to 1000nm wavelengths and with a sampling interval of 1.4 nm, and two indium gallium arsenide photodiodes covering 1000 to 2500nm wavelengths and with a sampling interval on 2 nm. The change between the two latter sensors occurs at 1830 nm. The data are interpolated by the spectrometers to 1nm intervals. The spectral resolution is of 3 nm at 700 nm, 8.5 nm at 1400 nm, and 6.5 nm at 2100nm (ASD, 2009).

FieldSpec 3Hi-Res was equipped with an ASD contact probe, designed to stay in a direct contact with solid materials and to reduce the errors related to the light source.

The spectra were acquired in the following conditions (fig. 5.1):

- measurements were carried out on 2 mm fine fraction of soil dried at air;
- the spectra were acquired in laboratory: samples were poured into a glass dish located on a black background;
- Spectralon® was adopted as white surface, recorded every 10 measurements for the calibration of the instrument; white reference measurements are needed to convert the measured radiance to relative reflectance;
- spectral signatures was created as mean of 20 measures with a duration of 0.1 s for each sample. For each sample three spectra were acquired at 3 different points of the sample, and the average value was recorded;
- reflectance was expressed as the ratio between the intensity of the light reflected from the reference material and from the soil sample.
- no preprocessing was applied to the data except for splice correction. This last provide a correction of the steps in an input spectral matrix by linear interpolation of the values of the edges of the middle sensor. The signatures with splice correction in the library of this study hereinafter are referred to as row spectra.

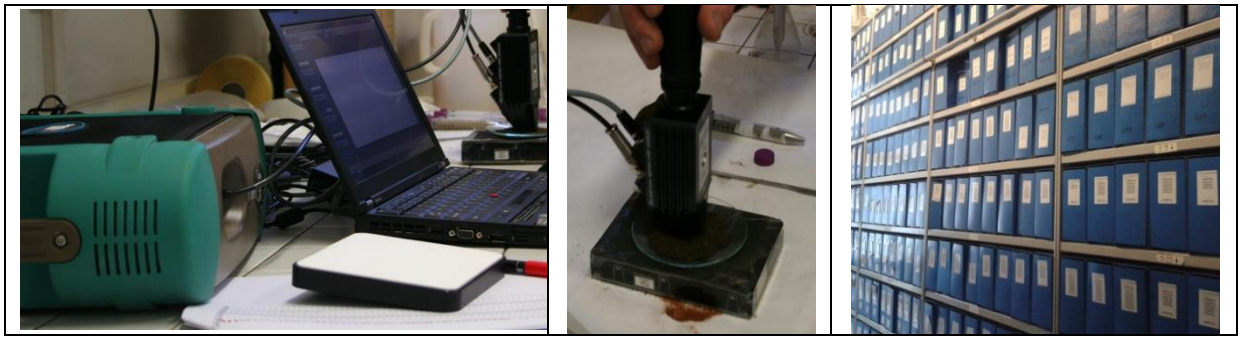


Figure 5.1. FieldSpec 3Hi-Res equipment; soil sample measurement; pedoteca of CREA.

5.2.4 Library organization

The library contains two main parts, the spectra files and the metadata. The metadata are composed of the available information collected into SISI database, a relational database in access (fig. 5.2). A table matches the name of the spectra signature file with the described and analyzed sample, while the spectra are stored separately from the database. Each soil spectrum is stored as single record into an all-inclusive ASCII file, with the mandatory name to be unique.

5.2.5. Italian spectral variability from free large library

In order to investigate over the usefulness of an Italian national library with respect to the data available for free at wider scale, the spectral library LUCAS was been queried. Indeed,

LUCAS covers all the European countries, including Italian territory. It contains geographic coordinates and information, and several analytical data. Moreover, data about soil forming factors and soil taxonomy are also available. Data located in the Italian territory were selected and a descriptive statistic was carried out on their metadata. The spectral variability and the pedological characterization of the Italian territory from LUCAS were compared with those obtained from the Italian National Library. Spectra collected in LUCAS were measured on air-dried and sieved (<2 mm) soil samples, with a XDS Rapid Content Analyzer (FOSS NIRSystems Inc., Laurel, MD). The spectrometer was equipped with Si (400–1100 nm) and PbS (1100–2500 nm) detectors, offering 4,200 wavelengths in the Vis-NIR region of the electro-magnetic spectrum (Stevens et al. 2013).

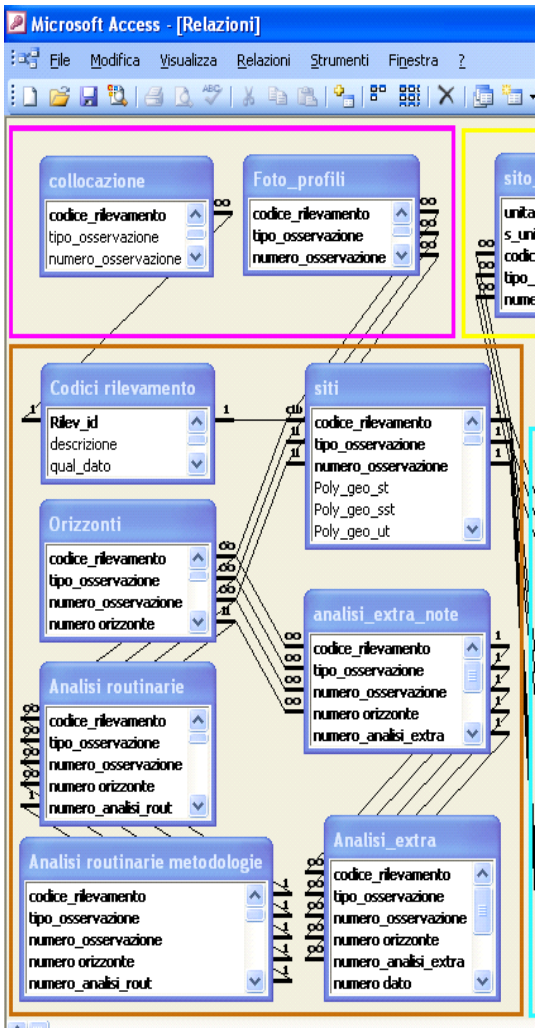


Figure 5.2: The Italian national soil database, SISI. The map reported the relations which can be used to associate the metadata to the spectral signature. Separated tables collect data about: the site, the field information on the horizons, the routinely analytical analyses, the additional analyses and the methods adopted.

5.2.6. Calibration and validation

5.2.6.1 Pre-treatment of spectral data

Mathematical and statistical procedures were carried out in Unscrambler (CAMO Software AS, Nedre Vollgate, Oslo, Norway).

In order to remove the parts of the spectra with low signal-to-noise ratios, wavelengths below 450 nm and above 2440 nm was discarded. This is in line with recommendations found in the literature: Seiler (2006) determined the signal-to-noise ratio for the used FieldSpec Pro device and recommended to discard wavelengths below 430nm and above 2440 nm.

Then, Savitzky-Golay smoothing filter (Savitzky & Golay, 1964) (2nd order polynomial covering 5 adjacent bands) was applied in order to reduce the noise. The pre-treated reflectance data was used for explorative data analysis reported in paragraph 2.6.2.

Alternatively, the raw data reflectance was transformed by derivatives of reflectance data, applying Savitzky-Golay filters of length 25 instead of the smoothing filter. The derivatives were used to reduce physical effect on the signature (e,g, scattering in the granular sample). The derivative transformation minimizes the effect of variation in sample grinding and optical set-up (Shepherd and Walsh, 2002).

5.2.6.2. Explorative analysis

A principal component analysis (PCA) was carried out as explorative investigation of the spectral variability in the Italian Spectral Library. PCA was carried out on data after smoothing by Savitzky & Golay.

Screening scatter plots of the scores of the first principle components was used to detect potential sub-groups within the dataset.

5.2.6.3 Calibration of models

Partial least square regression (PLSR) on mean centered data was used for modeling over the whole spectral dataset. First derivative of reflectance spectra was carried out in order to increase prediction accuracy, following the methodology explained in the paragraph *2.6.1*.

Square root transformation of the laboratory measurements was also adopted for estimating SOC as it was evaluated to improve the results.

Cross-validation was used for determining model architecture and prediction testing for testing the final performance of the model.

Models were evaluated in two ways: (1) leave-one-out cross-validation on models developed with all the available samples and (2) 10-fold cross-validation by splitting the sampling set into spectral calibration (90% of the dataset) and validation (10% of the dataset) sets.

The number of factors to retain in the calibration model was determined by means of the 10-fold cross-validation.

Data which were considered outliers was excluded from the prediction in order to outperform both the SEP and the BIAS of the models. SEP is defined as the standard deviation of the predicted residuals, thus it is a measure of the precision; BIAS is the average difference

between observed and predicted values in the prediction set. The samples were removed from the calibration when they resulted in an extreme position in the population, or when the model produced a very poor prediction.

Prediction accuracy was assessed by root mean squared error (RMSE) of the cross-validation ($RMSECV_{10}$ for the 10-fold cross-validation; $RMSECV_f$, for the leave one out cross-validation), by the coefficient of determination R^2, and the ratio of standard deviation (SD) to $RMSECV_{10}$ (RPD_{10}) and (Chang et al., 2001; Waiser et al., 2007).

5.3. Results and discussion

5.3.1 Metadata: representativeness of Italian soil variability by Italian Spectral Library and a comparison with LUCAS

Currently the spectra signatures are available for 1179 samples of Italian soils, collected from 715 soil profiles, and distributed in 874 top soil (upper limit lower than 30 cm) and 305 subsoils (upper limit deeper than 30 cm). Table 5.1 highlights a different availability of metadata in the library. Organic Carbon and total carbonates are the most available data, followed by texture, AWC, total Nitrogen and pH in water. Active carbonate, ESP, electrical conductivity and bulk density have a lower frequency. According to the varied origin of the sampled collected, the ranges of the soil attributes result wide and their coefficient of standard deviation large.

The most represented taxonomic groups were the following Reference Soil Groups of the WRB: Cambisols (290 samples), Calcisols (119), and Luvisols (75), followed by Regosols Pheozems and Vertisols with about 50 samples, and Andisols Kastanozems and Alisols with about 30 samples. This stratification partially agreed with the Italian Coverage of dominant Reference Soil Group of WRB. As reported in the book Italian soil of Italy (Costantini et al., 2013), most of the Italian soils belong to Cambisols (39% of the surface), Luvisols (13%), Regolos (10%), Phaeozem (8%), and Calcisols (8%). While, in the legend of the map of the soil system of Italy, the most common kind of soils are Haplic Cambisols (Calcaric), followed by Haplic Regosols (Calcaric), Cambisols (Eutric), Haplic Calcisols, Vertic Cambisols, Cutanic Luvisols, Leptic Phaeozems, Haplic Luvisols (Chromic), Haplic Cambisols (Dystric), and Fluvic Cambisols (tab. 5.2).

Table 5.1: Main metadata of the Italian spectral library.

pamater	n hor.	Mean	SD	Max	Min
Clay (%)	865	34.7	15.5	93	0.8
Silt (%)	863	38.8	21.1	92.7	0.2
Sand (%)	863	26.4	21.4	98.5	0.6
OC (%)	1179	2.1	2.8	33.4	0
N (%)	841	1	1.1	14.6	0
$CaCO_3$ tot (%)	1057	10.7	15.3	78.3	0

CaCO3 active (%)	417	3.2	3.5	24.4	0
pH	676	7.7	0.9	9.4	2.8
CEC cmol(+)/kg	540	21.5	16.3	100	1.4
Ca cmol(+)/kg	599	19.6	17.7	257	0.1
Mg cmol(+)/kg	598	2.6	2.4	17.8	0
Na cmol(+)/kg	601	0.9	1.7	15.4	0
K cmol(+)/kg	601	0.5	0.5	5.5	0
EC (dS/m)	567	0.4	0.6	3.9	0
ESP (%)	571	5.6	8.1	68.7	0
AWC (mm/m)	852	144.6	27.8	233.6	54
BD (g/cm3)	487	1.2	0.3	2.1	0.2

Table 5.2: Distribution of spectral signatures of the Italian spectral library across the WRB orders (CM: Cambisol; CL: Calcisols; LV: Luvisols; RG: Regosols; PH: Phaeozems; VR: Vertisol; AN: Andosols; KS: Kastanozems; AL: Alisols; LP: leptosols; GL: Gleysols; CH: Chernozems; FL: Fluvisols; LX: Lixisols; AR: Arenosol; AB: Albeluvisol; NT: Nitisol; ST: Stagnosol; SN: Solonetz).

WRB	CM	CL	LV	RG	PH	VR	AN	KS	AL	LP	GL	CH	FL	LX	AR	AB	NT	ST	SN	Null
N	290	119	75	53	51	51	31	29	26	23	21	15	15	6	6	6	3	3	2	354

The Italian samples stored in LUCAS achieved the number of 1180 spectral signatures distributed over all the Italian territory. As highlighted from the maps in the figure 5.3, the samples showed a more homogeneous distribution in LUCAS datased than in the Italian library at the state of the art. However, the range of pedological variability covered by LUCAS is lower than the variability associated to the signatures in the Italian spectral library (table 5.3). Also the taxonomic pedodiversity in LUCAS is lower: only 7 reference soil groups are represented, respect to the 17 than the Italian library (table 5.4).

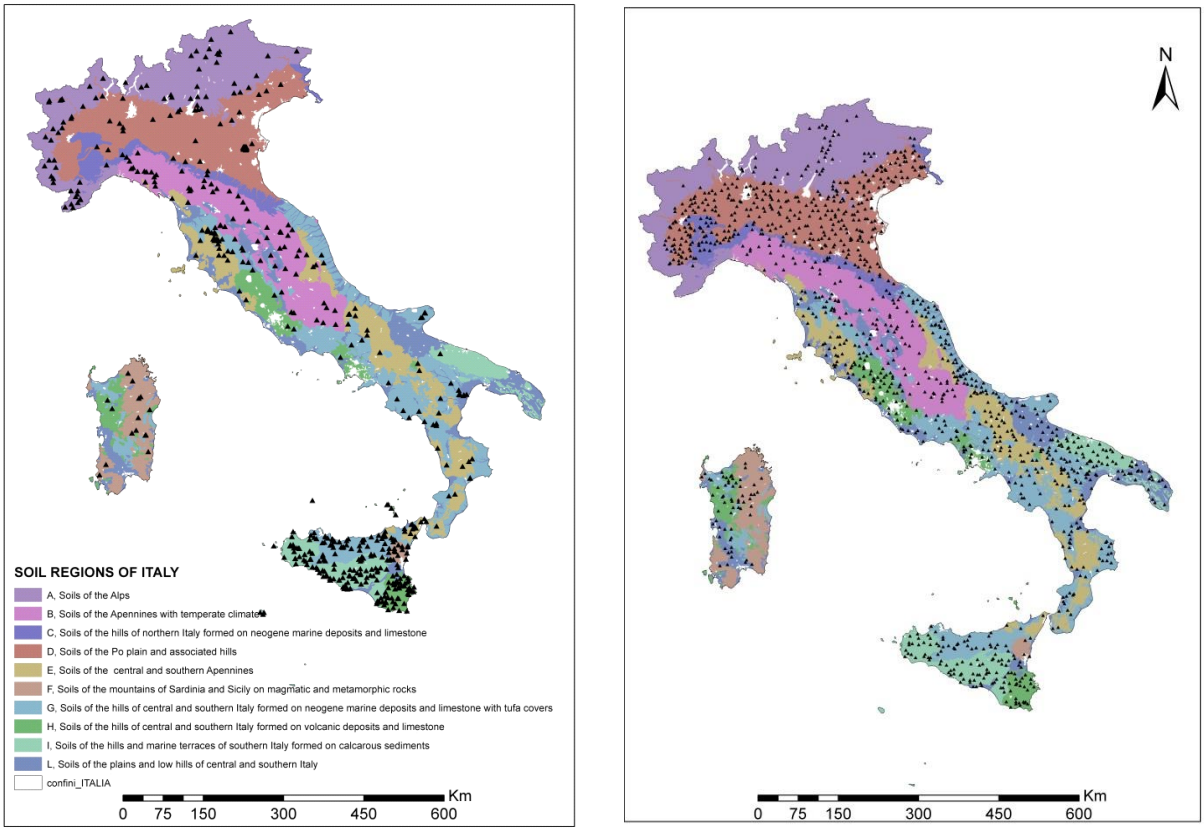


Figure 5.3. Spatial distribution of spectral information over the Italian territory according to the Italian spectral library (on the left) and LUCAS library (on the right).

Table 5.3: Main metadata of LUCAS library for the Italian territory

	n hor.	Mean	SD	Max	Min
Clay (%)	1180	29.4	8.49	79	2
Silt (%)	1180	42.8	2.12	80	3
Sand (%)	1180	27.9	11.3	92	1
OC (%)	1180	2.11	1.65	23.77	0
N (%)	1180	0.19	0.06	1.57	0
$CaCO_{3\ tot}$(%)	1180	11.65	24.4	76.2	0
pH	1180	7.4	1.1	8.8	3.9
CECcmol(+)/kg	1180	22.7	0.99	93.1	2.2

Table 5.4: Distribution of spectral signatures of LUCAS library for the Italian territory across the WRB orders (CM : Cambisols; LV: Luvisols ; RG: Regosols; FL: Fluvisols; VR: Vertisol; LP: leptosols; AN: Andosols).

WRB	CM	LV	RG	FL	VR	LP	AN	Null
N	958	89	43	29	24	19	14	4

5.3.2 Spectral variability

The spectra measurements of the samples collected in the Italian library are shown in figure 5.4. The spectra demonstrated the basic variability in the soil spectral shapes across the VisNIR (400–2500 nm) wavelength ranges. In the figure, the raw data only were preprocessed by splice correction. The raw spectra generally showed the typical pattern of soil spectra, with three major absorption features. The first absorption region near 1400 nm corresponds to the first overtone of OH stretches (related to the water adsorbed to the clay surfaces); at about 1900 nm a peak corresponds to the combination of OH stretches and H-O-H bend in water molecules included in the crystal lattice (as for example in illite and montmorillite). Near 2200 nm it is OH-metal bend and OH stretch combinations, it could be due to Al or Fe or Mg substituting Si (Clark et al., 1990).

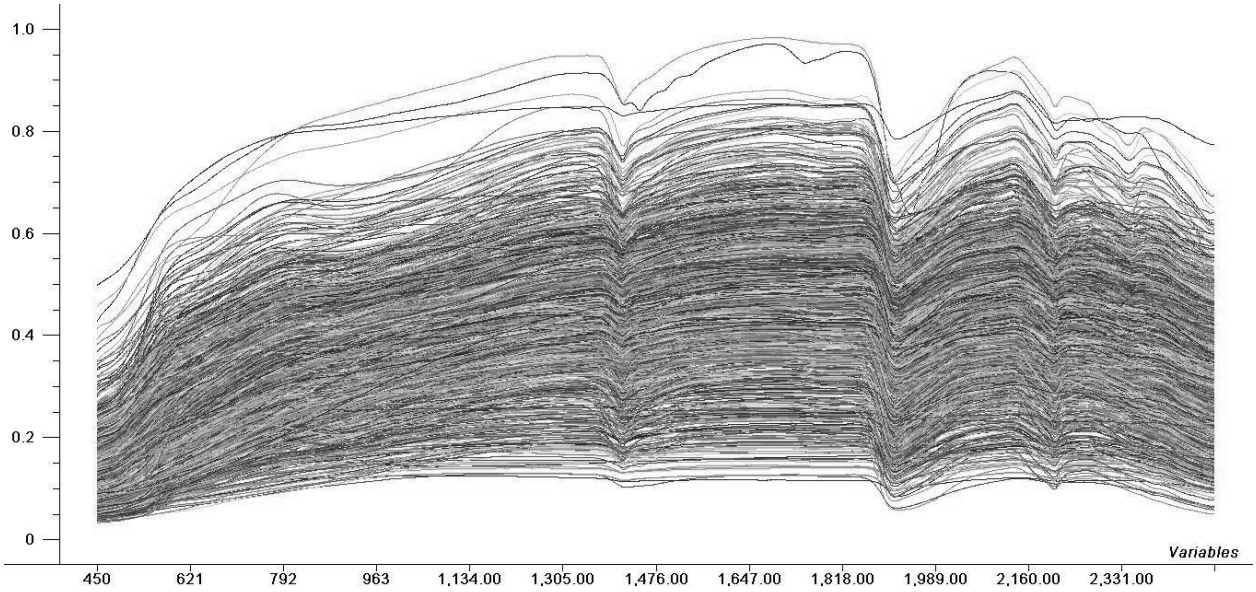


Figure 5.4: Spectra of the soil samples showing the major absoption features, related to OH groups in both absorbed water (about 1400 and 1900 nm) and the crystal latticine (about 2200 nm).

Results for the principal component analysis (PCA) showed that the first three principal components (PCs) accounted for 94% of the total variation. Figure 5.5 shows scatter score plots of the first three PCs with the corresponding loading line plots. The PC scores in the scatter plots were grouped into five spectral groups which correspond to the number of clusters automatically chosen by the general K-means classification, based on PCs score. Sample grouping for PC1 vs. PC2 plot clearly split each form the other the spectral cluster 1, 2, 3, and 5. Cluster 4 was better separated by PC1 vs. PC3.

The loading of PC1 shows mean peaks at around 1400 and 1900 nm and 2200 and 2300nm, which could be assigned respectively to absorbed water and to the presence of clay minerals such as kaolinite, montmorillonite. and illite due to the combination of vibrations

associated with the OH bond and the OH-Al-OH bonds (Hunt et al.,1977 Chabrillat et al., 2002) PC2 accounts for 3% of the total variation. It shows a significant peak around 580 nm, which could be assigned to OM. As the first PC there are peaks at around 1400 and 1900 nm and 2200. PC3 accounts for 2% of the variation, it has pronounced signals in the visible range (about 470 nm) that could be assigned to Fe oxides, whereas peaks near 1400 and 1900 nm and 2200 nm are in common with the other PCs .

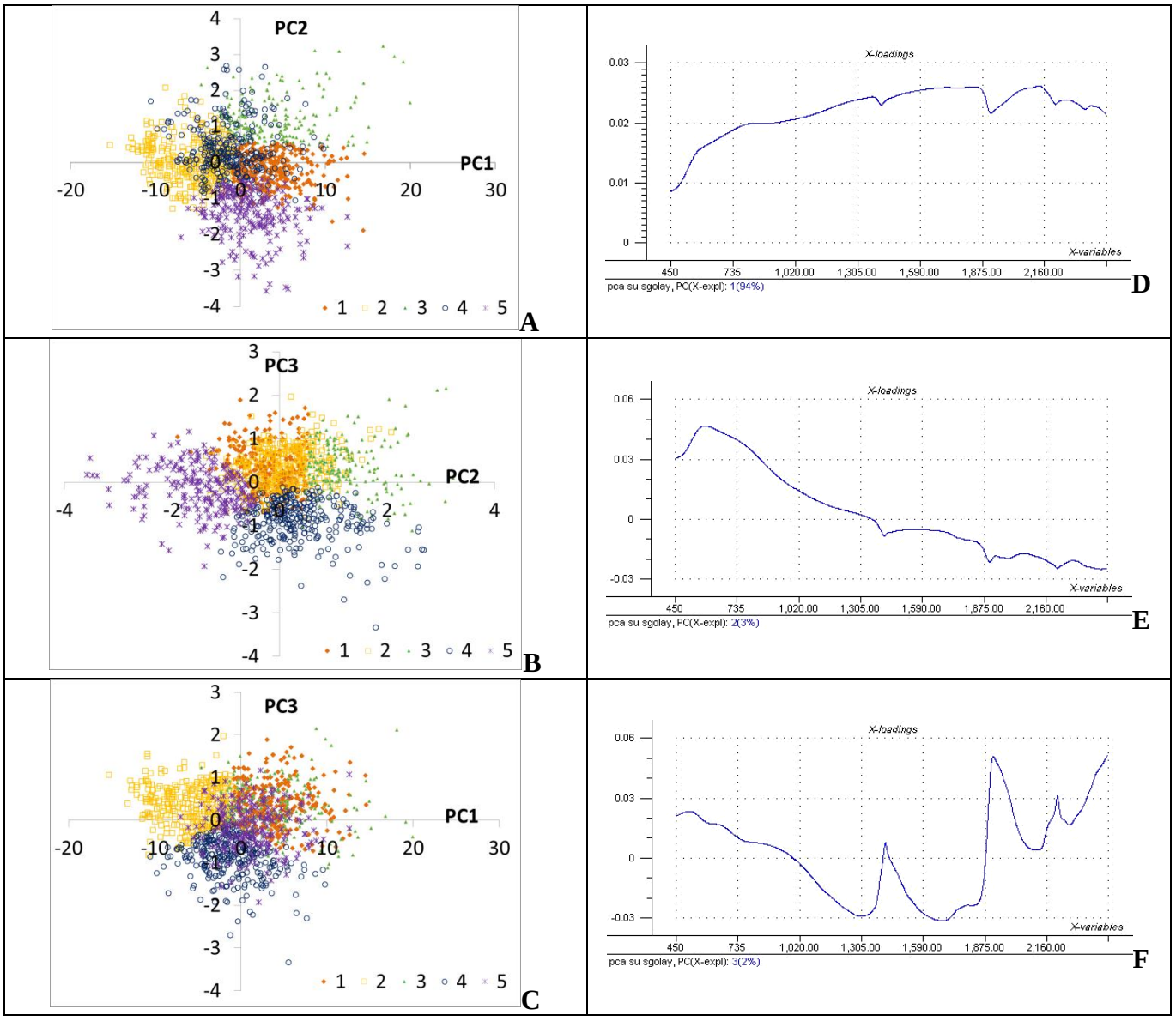


Figure 5.5: PC score plots with sampling spectral groups (A:PC1-PC2, B:PC2-PC3, C:PC1-PC3) and loading plots for the entire data set (E: PC1, F: PC2, G:PC3)

In order to carry out an explorative investigation, the PCs scores were crossed with some soil properties known to be well correlated with spectral information. Organic carbon, texture, and carbonates were separately used to classify the samples in subjective groups with increasing

values of the variables. Actually, the plots did not highlight a clear distribution of the clusters along any of the three PCs. It is only possible to appreciate: the group of fine texture (clay >40%) towards negative values of PC1 and PC2, while the other texture classes are more concentrated on the central value of the PCs; the group of very high OC content (>10%) towards negative values of PC1 (figg. 5.6, 5.7, 5.8)

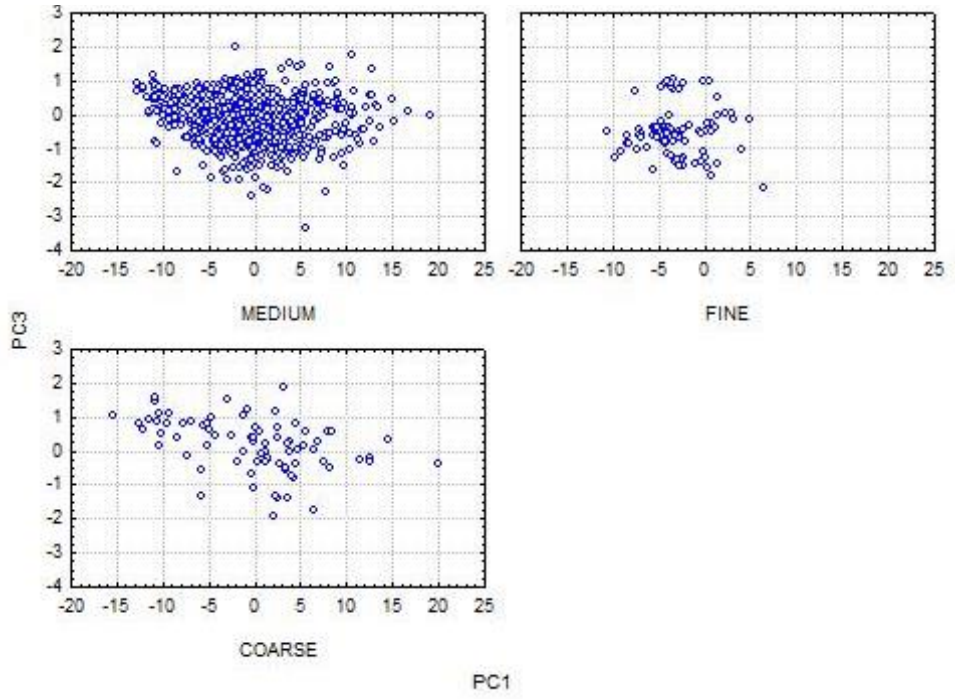


Figure 5.6: PC1-PC2 score plots with sampling grouping according to texture (fine: clay>40%; coarse: sand> 60%; medium: all the other

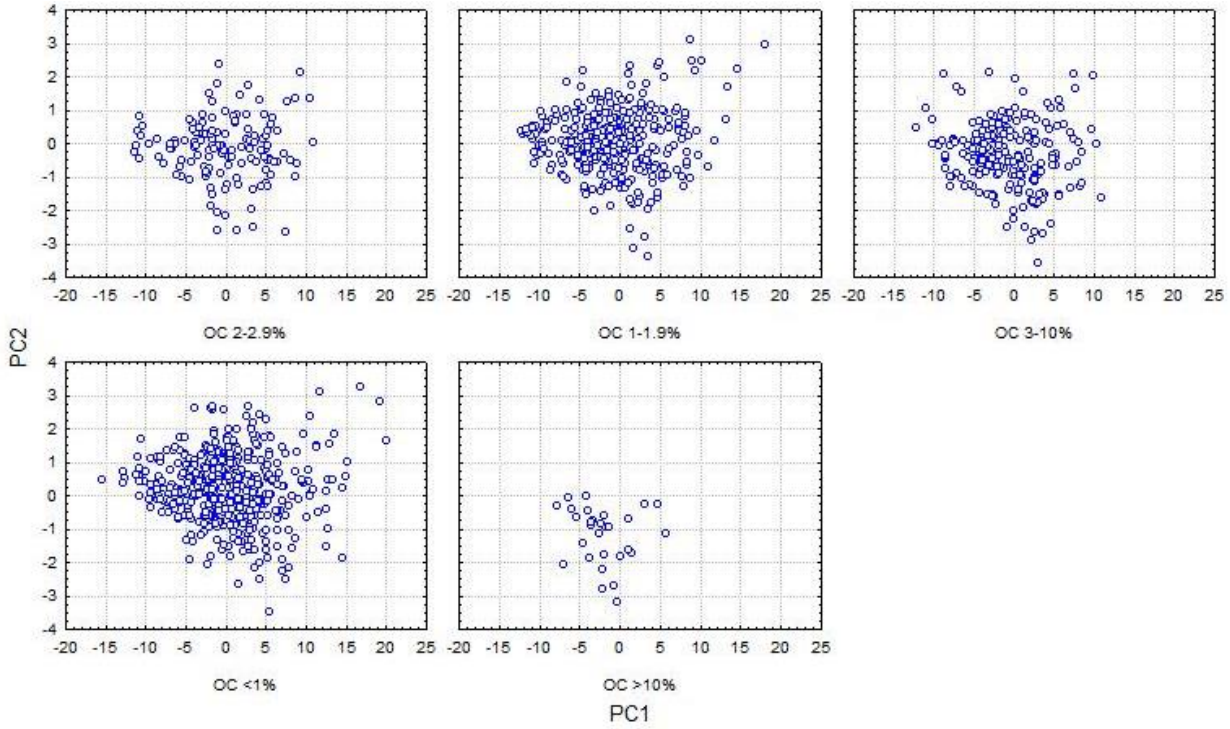


Figure 5.7: PC1-PC2 score plots with sampling grouping according to SOC content (%)

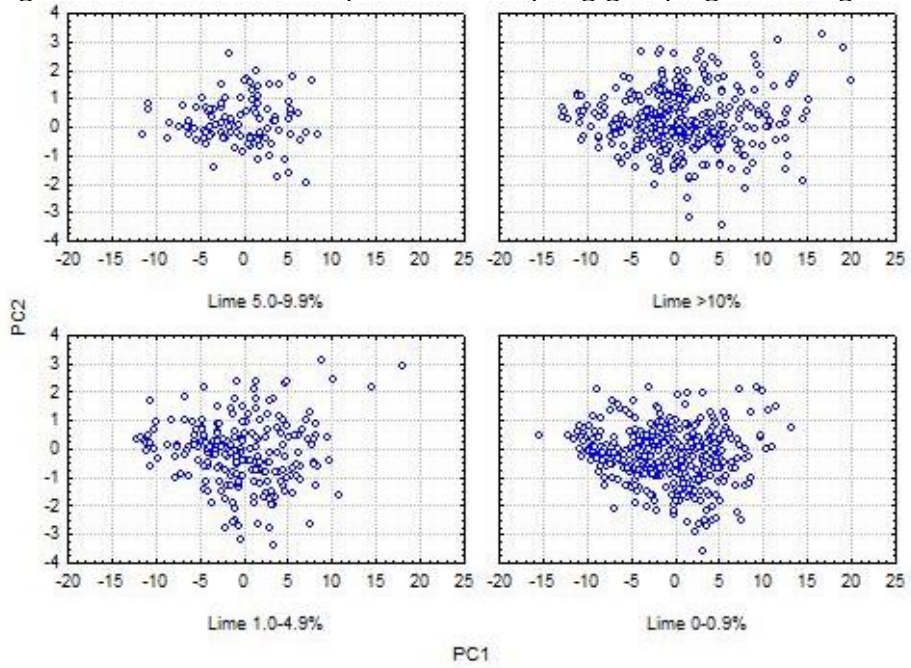


Figure 5.8: PC1-PC2 score plots with sampling grouping according to total carbonate content (%)

5.3.3 Example modeling

Some models were calibrated on the whole dataset. The OC, clay and carbonates contents were selected, since they are known to have a good predictability by VisNIR.

In this study the error in prediction ($RMSECV_{10}$) showed to be affected mainly by SEP_{10}, thus the prediction was not biased, but it showed lack of precision.

Due to the very high variability of those variables, national models were not expected to be very accurate. Indeed, the $RMSECV_{10}$ (table 5.5) was rather high in all the cases if compared with the mean values (see table 5.3). However, the results showed a fair predictability with RPD around 2. The good accuracy indicated by RPD depends on the fact that very large standard deviation occurred in the dataset even when the outliers were kept out of the calibrations. In another study carried out at national scale on a very large amount of Danish soils (2851 samples), PLSR for SOC estimation gave similar R^2 for the validation (0.78), a little better RPD (2.3) and a clearly better RMSE of prediction (0.31) (Knadel et al., 2012). In this comparison it is important to note that samples in the Italian national library cover a very larger range of values (SD of 1.67 vs 0.71), so that a little poorer SOC prediction was expected to be obtained. In the same way, a prediction of SOC content carried out on a national VisNIR spectroscopic library in Florida by means of 7120 samples leaded to a similar RPD (2.14) but to a very high RMSE of validation (2.52), that can be attributed to the very high standard deviation of the dataset (5.17) (Vasques et al., 2010). The same concept emerged by a study over the clay content predicting by means of a National VisNIR library of soils from Brazil (Araújo et al., 2014). Their finding was a prediction with RMSE of 10.9 and an RPD of 2.4 when PLSR was carried out over the database as a whole (SD=26%). Then, their global data set was clustered into groups which were of uniform mineralogy, regardless of geographical origin, by dividing it into smaller sub-libraries on the basis of the vis-NIR spectra. The clustering improved predictive performance and the RMSE was enhanced of 8.6%.

Table 5.5: Partial least square regression results for SOC (back-transformed), clay, and total carbonate content, (using 1'st derivate Savistzky Golay derivate). Parameters referred to predicted values by 10-fold cross-validation; *: SD of the reference values.

	SOC (%)	Clay (%)	Total $CaCO_3$ (%)
Original n. of samples	1182	865	1057
n. samples in the model	1064	749	810
n. predictive factors	7	7	6
n. of outliers	118	116	247
R^2	0.753	0.774	0.872
$RMSECV_{10}$	0.842	7.69	4.76
SEP_{10}	0.840	7.69	4.76
$BIAS_{10}$	-0.071	0.001	-0.002
SD *	1.67	16.2	13.3
RPD	1.98	2.10	2.79

5.3.3.1 Calibrating model for SOC prediction

Looking at the data excluded from the calibrations (fig. 5.9), they were those at the border of the population. Actually, the population used for the model excluded values above 15, while the population of excluded data showed a higher frequency at the extreme values of OC (OC = 0 and OC> 15). The samples out of the elaboration belonged mostly to forestall stations and soils with gypsum (table 5.6). Keeping them out of the calibration, the overall variance was reduced, and the prediction improved.

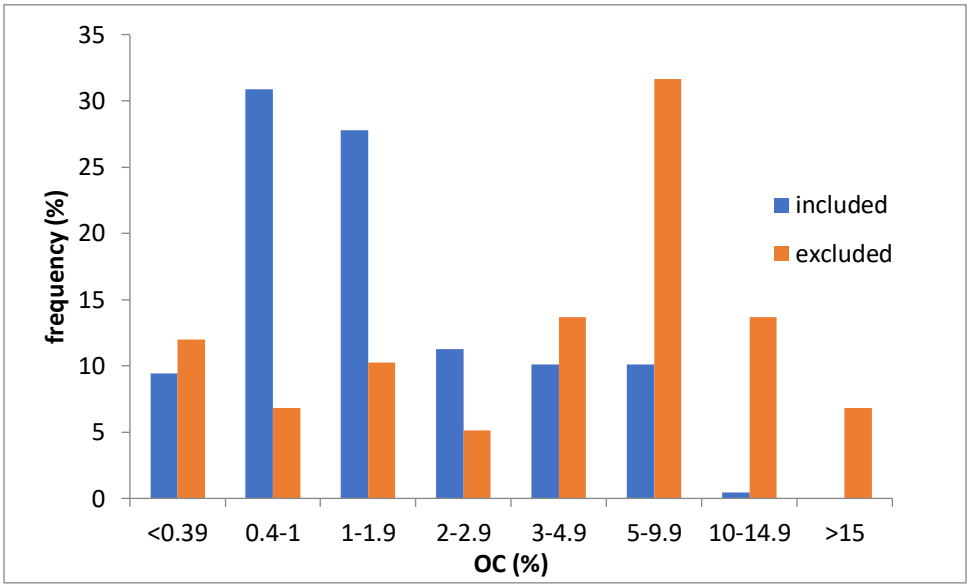


Figure 5.9: Distribution of samples according to the OC content (%) in the model (A) and in the outliers (%) for the prediction OC content at national scale.

Table 5.6: Descriptive Statistics (PC.smx) Excluded against employed

	Valid N	Mean	Minimum	Maximum	Std.Dev.	% forestall soil	% soil with gypsum
employed	1065	1.67	0.00	15.2	1.68	437 (41%)	69(6.4%)
Excluded	117	6.16	0.00	33.4	5.65	72 (61.5%)	11 (9.4%)

5.3.3.2 Comparing 10-fold and leave-one-out cross-validation

Segmented cross-validation (namely 10-fold) may be very useful if some structures exist in the dataset, for instance, when samples are clustered. Clustering may occur when samples are collected in a little area under the effect of the same pedogenetic factors, or they belong to the same profile. In these cases, leave-one-out cross-validation may give over-optimistic results. On the other hand, the leave-one-out cross-validation is more significant in situations with randomly selected calibration samples from a natural population (Naes et at., 2002). Since this library was created by means of samples collected along the years for different purpose, both in very

restricted and wider areas, a clustering may occur across the data. So, we evaluated the models on the base of 10-fold cross-validation. However, a leave-one-out cross-validation to compare the results and over-optimistic result was expected. Conversely, the cross-validation showed the same efficiency of the predictive models (table 5.7) It could indicate that, despite the origin of the data, the library was not so affected by clusterization to have a relevant effect on the predictions.

Table 5.7: Comparing PLSR efficiency parameters between 10-fold ($*_{10}$) and leave-one-out ($*_f$) cross-validation.

PLSR parameters	$RMSECV_{10}$	$RMSECV_f$	SEP_{10}	SEP_f	$BIAS_{10}$	BIASf	$Correlation_{10}$	$Correlation_f$
OC (%)	0.84	0.084	0.84	0.84	-0.071	0.074	0.87	0.87
Clay(%)	7.69	7.31	7.69	7.32	0.001	0.006	0.879	0.892
CaCO3 (%)	4.76	4.6	4.76	4.76	-0.002	-0.006	0.933	0.933

5.4 Conclusions

This first version of the Italian spectral library contains over 1179 collected spectra. The version includes the basic soil spectra variation of the Italian soils. The pedological variability covered by the collected samples is higher that those covered by LUCAS for the Italian territory. It testifies the importance that still may have the development of a national database, since one of the main topics of a wide library is covering most of the variability. In this way, it may allow predictions in a wider range of values of the target variables. Moreover, the possibility to extract a part of the library for prediction over a specific typology of soil is bigger when a bigger number of soil typology are represented in the library.

The constantly updated library provides an initial collection of spectra covering the VisNIR wavelength range. It allows for modelling soil properties by statistical inference. Here, PLSR models for organic carbon, clay and carbonates were carried out with good results, supporting that it is suitable for further soil spectroscopy exploration in the Italian territory.

Acnoledgments

I wish to thank Simone Priori and Giovanni L'Abate for their help in collecting and elaborating data.

References

Araújo, S. R., Wetterlind, J., Demattê, J. A. M., Stenberg, B., 2014. Improving the prediction performance of a large tropical vis-NIR spectroscopic soil library from B razil by

clustering into smaller subsets or use of data mining calibration techniques. European Journal of Soil Science, 65(5), 718-729.

Baumgardner, M.F., Silva, L.F., Biehl, L.L., Stoner, E.R., 1985. Reflectance properties of soils. Adv. Agron. 38:1–44.

Bellinaso, H., Demattе, J.A.M., Araujo, S.R., 2010. Spectral library and its use in soil classification. Braz. J. Soil Sci. 34, 861-870.

Ben-Dor, E., Banin, A., 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Science Society of American Journal 59, 364–372.

Brodský, L., Klement, A., Penížek, V., Kodešová, R., Borůvka, L., 2011. Building soil spectral library of the Czech soils for quantitative digital soil mapping. Soil Water Res, 6(4), 165-172.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132, 273–290.

Chabrillat, S., Goetz, A. F., Krosley, L., Olsen, H. W., 2002. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. Remote sensing of Environment, 82(2-3), 431-445.

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh Jr., C.R., 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. Soil Science Society of America Journal 65 (2), 480–490.

Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. Journal of Geophysical Research 95 (B8),

Costantini, E. A., Barbetti, R., Fantappie, M., L'Abate, G., Lorenzetti, R., Magini, S., 2013. Pedodiversity. In The soils of Italy (pp. 105-178). Springer, Dordrecht.

Dalal, R.C., Henry, R.J., 1986. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. Soil Sci. Soc. Am. J. 50, 120–123.

Demattê, J.A.M., Campos, R.C., Alves, M.C., Fiorio, P.R., Nanni, M.R., 2004. Visible–NIR reflectance: a new approach on soil evaluation. Geoderma 121, 95–112.

Genot, V., Colinet, G., Bock, L., Vanvyve, D., Reusen, Y., Dardenne, P., 2011. Near infrared reflectance spectroscopy for estimating soil characteristics valuable in the diagnosis of soil fertility. J. Near Infrared Spectrosc. 19 (2), 117–138.

Gogé, F., Joffre, R., Jolivet, C., Ross, I., Ranjard, L., 2012. Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database. Chemometrics and Intelligent Laboratory Systems,110(1), 168-176.

Knadel, M., Deng, F., Thomsen, A., Greve, M. H., 2012. Development of a Danishnational vis-NIR soil spectral library for soil organic carbon determination. In Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping 2012, Sydney, Australia (Vol. 403). CRC Press.

Hunt, G. R. 1977. Spectral sigantures of particulate minerals in the visible and near infrared. Geophysics 42:501-513.

Ji,W., Viscarra Rossel, R.A., Shi, Z., 2015. Accounting for the effects of water and the environment on proximally sensed Vis–NIR soil spectra and their calibrations. Eur. J. Soil Sci. 66 (3), 555–565.

Næs, T., Isaksson, T., Fearn, T., Davies, T., 2002. A user friendly guide to multivariate calibration and classification. NIR publications.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., ..&. Dardenne, P. (2015a). Soil spectroscopy: An alternative to wet chemistry for soil monitoring. In Advances in agronomy (Vol. 132, pp. 139-159). Academic Press.

Nocita, M., Stevens, A., vanWesemael, B., Brown, D.J., Shepherd, K.D., Towett, E., Vargas, R., Montanarella, L., 2015. Soil spectroscopy: an opportunity to be seized. Glob. Chang. Biol. 21 (1), 10–11.

Reeves III, J.B., McCarty, G.W., 2001. Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and fibre-optic probe. Journal of Near Infrared Spectroscopy 9, 25–34.

Sanchez, P.A., Ahamed, S., Carré, F., Hartemink, A.E., Hempel, J., Huising, J., Lagacherie, P., McBratney, A.B., McKenzie, N.J., Mendonça-Santos, M.d.L., Minasny, B., Montanarella, L., Okoth, P., Palm, C.A., Sachs, J.D., Shepherd, K.D., Vågen, T.-G., Vanlauwe, B., Walsh, M.G., Winowiecki, L.A., Zhang, G.-L., 2009. Digital soil map of the world. Science 325 (5941), 680–681.

Savitzky, A., Golay, M. J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry, 36(8), 1627-1639.

Shepherd, K.D.,Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66, 988–998

Stenberg, B., Nordkvist, E., Salomonsson, L.,1995. Use of near infrared reflectance spectra of soils for objective selection of samples. Soil Science 159, 109–114.

Stevens, A., Nocita, M., Tóth, G., Montanarella, L., & van Wesemael, B. (2013). Prediction of soil organic carbon at the European scale by visible and near infrared reflectance spectroscopy. PloS one, 8(6), e66409.

Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. Soil Science Society of America Journal 71 (2), 389–396.

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(1), 29-38.

Vasques, G. M., Grunwald, S., Harris, W. G., 2010. Spectroscopic models of soil organic carbon in Florida, USA. Journal of environmental quality, 39(3), 923-934.

Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158 (1–2), 46–54.

Viscarra Rossel, R. V., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D.,. Shi Z., Stenberg B., Stevens A., Adamchuk V., Aïchi H., Barthès B.G., Bartholomeus H.M., Bayer A.D., Bernoux M., Böttcher K., Brodský L., Du C.W., Chappell A., Fouad Y., Genot V., Gomez C., Grunwald S., Gubler A., Guerrero C., Hedley C.B., Knadel M., Morrás H.J.M., Nocita M., Ramirez-Lopez L., Roudier P., Rufasto Campos E.M., Sanborn P., Sellitto V.M., Sudduth K.A., Rawlins B.G., Walter C., Winowiecki L.A., Hong S.Y., Ji W., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews, 155, 198-230.

Viscarra Rossel, R.A., Lark, R.M., 2009. Improved analysis and modelling of soil diffuse reflectance spectra using wavelets. Eur. J. Soil Sci. 60 (3), 453–464.

Viscarra Rossel, R.A., Walvoort, D.J.J., 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 (1–2), 59–75.

Viscarra Rossel, R.A., Webster, R., 2012. Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. Eur. J. Soil Sci. 63 (6), 848–860.

CAP 6.

GENERAL CONCLUSIONS

Regarding the general issue of this thesis it can be concluded that it has been possible to approach with success several different topics of the soil science by using spectroscopic information.

Each of the research chapters in this thesis has shown that it is possible to effectively describe by spectral means the soil variation contained at different levels and observed by different points of view of soil investigation. The results try to fill out the idea of the potential of using VisNIR spectroscopy to predict soil properties and discriminate soil types or soil samples. The importance of this potential is actually increasing day by day since soils are crucial global issues as involved in several international policies for food production, climate change and environmental protection.

In particular, the first approached subject deals with the importance to find a smart way for increasing the basic knowledge of soils. Recognizing and ordering soils in tassonomic clusters is at the base of this fundamental knowledge and it is an indispensable support for the right management of the soil resource, since each soil typology has its capatilities and points of criticism. The investigation in the frame of soil taxonomy presented in the chapter 2, showed the close connection between soil spectroscopy attribute variation and soil taxonomic units linked to volcanic properties. The thesis introduced at the possibility to employ VisNIR information for a taxonomic classification of soils with andic properties by means of predictive calibrations and classifications. Models could be used as a preliminary easy method to distinguish Andosols from other soils and to highlight different degrees of andic properties. On the base of the encouraging results achieved with some calibration models (1.5>RPD>1.9), further wider studies, could lead to more reliable models and return quantitatively accurate estimations for several investigated andic parameters.

In respect to the topic of soil monitoring, the results of the chapter 3 gave an indication of the necessary conditions for providing sufficient advantage of VisNIR in monitoring SOC, such as to outperform traditional wet chemistry. Viscarra Rossel et al. (2016), remarked on the need for more research to optimally use large spectroscopic databases for local predictions of soil attributes. With respect of this, extra-weighted spiked models demonstrated to be the most appropriated for comparing SOC by means of NIR. It gave the lowest BIAS, and therefore, the lowest error in reproducing real differences between the compared data. We can assume an economic benefit with NIR spectroscopy, when small changes should be detected, as those expected in less than 10 years. Moreover, it is important to consider that, if monitoring activity provides repeated assessments, the cost will favour the use of NIR, since there is no need to analyse the spiking subset (i.e., the model does not need to be adapted each time) and those initial efforts invested on adapting the models might be less important in the overall budget. Thus, after the repetition of several assessments (successive events), the NIR is expected to surpass WB in a higher number of cases. Finally, if the VisNIR measurements are carried out in

laboratory, like in our study, the main cost of monitoring studies by VisNIR are linked to field campaigns. Therefore, an overmore increasing of cost benefit may occur by the use of models based on portable VisNIR instruments. Actually, Vis-NIR spectroscopy is currently mostly used in laboratory conditions, but its application in-situ and even on air- or space-borne platforms is growing (Ben-Dor et al., 2009). Further studies may focus on a quantification of these two last aspects for a very exhaustive discussion over this topic.

Always in the frame of soil monitoring activity, the chapter 4 of this thesis was about the possibility to describe the soil variability directly by VisNIR properties. It was observed that the multivariate information contained in the soil spectrum was as hypothesized, useful in describing the variations occurring in the soil as a consequence of the variation occurring in the landscape. The results allowed to affirm how a very little variation induced by a recent change in the management of a forest landscape can be captured by VisNIR data better than biochemical information. According to its nature, VisNIR gave integrated information on physical, chemical, and biological feature of the sample. It can justify the higher power in detaching slight differences due to effect of a very recent changing in the forest management.

In the chapter 5, it is shown how, despite the presence of wide free spectral libraries, as other authors have concluded e.g., (Guerrero et al., 2016), the creation of large spectral libraries is still a challenge. Indeed, the creation of a national library allowed to better cover the pedological variability of our country. This thesis represents the first drafting of the Italian spectral library, and, at the state of the art, it is able to cover a higher pedological variability respect to free spectral library already available for the same territory.

As a whole, the case studies deal with the application of VisNIR spectroscopy: at different scale region (European, national, and local); on different kind of libraries (thematic as volcanic soils, and wide library both for local predictions and for national models); in different aspects of the pedological investigations, from recognizing taxonomic groups, to the ability to track changes in soil properties. Thus, it represents a wide-reaching contribute to clarify the usefulness of VisNIR spectroscopy for the major soil issues (e.g. quick data acquisition, soil C accounting, detection of changes in land uses and management practices).

References

Ben-Dor E, Chabrillat S, Demattê JAM, Taylor GR, Hill J, et al. (2009) Using Imaging Spectroscopy to study soil properties. Remote Sensing of Environment 113: S38–S55.

Guerrero, C., Wetterlind, J., Stenberg, B., Mouazen, A.M., Gabarrón-Galeote, M.A., Ruiz-Sinoga, J.D., Zornoza, R., Viscarra Rossel, R.A., 2016. Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? Soil and Tillage Research 155, 501-509.

Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aïchi, H., Barthès, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Böttcher, K., Brodský, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C.B., Knadel, M., Morrás, H.J.M., Nocita, M., Ramirez-Lopez, L., Roudier, P., Campos, E.M.R., Sanborn, P., Sellitto, V.M., Sudduth, K.A., Rawlins, B.G., Walter, C., Winowiecki, L.A., Hong, S.Y., Ji, W., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews.