Scientific synthesis of the DIGISOIL project

Gilles Grandjean, Sébastien Lambot, Isabelle I. Cousin, Bas van Wesemael, Iason Diafas, Leandro Chiarantini, Olivier Cerdan

Public document

Scientific synthesis of the DIGISOIL project

FP7 – DIGISOIL Project Deliverable D5.1

N° FP7-DIGISOIL-D5.1
September 2011

…

The DIGISOIL project (FP7-ENV-2007-1 N°211523) is financed by the European Commission under the 7th Framework Programme for Research and Technological Development, Area "Environment", Activity 6.3 "Environmental Technologies".

Geoscience for a sustainable Earth

Public document

Scientific synthesis of the DIGISOIL project

FP7 – DIGISOIL Project Deliverable D5.1

N° FP7-DIGISOIL-D5.1
September 2011

G.Grandjean (BRGM)
With the collaboration of
DIGISOIL Team

Checked by:	**Approved by**:
Name: S.Lambot	Name: A.Vagner
Date:	Date:
☑	☑

BRGM's quality management system is certified ISO 9001:2000 by AFAQ.

The DIGISOIL project (FP7-ENV-2007-1 N°211523) is financed by the European Commission under the 7th Framework Programme for Research and Technological Development, Area "Environment", Activity 6.3 "Environmental Technologies".

Geoscience for a sustainable Earth
brgm

Keywords: geophysics, sensors, soil properties

In bibliography, this report should be cited as follows:
G.Grandjean and DIGISOIL Team, 2011. Scientific synthesis of the DIGISOIL project. Report N° FP7-DIGISOIL-D5.1; 54 pages.

Synopsis

Since 2002, the European Commission considers the soil as a non-renewable resource threatened by natural hazards and human activities. Erosion, organic matter decline, contamination, sealing, compaction, salinisation, floods and landslides have been identified as the main soil threats. The implementation of the Soil Thematic Strategy (COM232, 2006) implies that Member States individuate risk areas for erosion, organic matter decline, compaction, salinisation and landslides. Among the main soil and soil-related parameters to be considered by Member States for delineating risk areas, soil texture, soil water content – and related hydraulic properties, bulk density and soil organic matter, impact crucially many soil functions, so that these properties have to be considered as a first priority. In particular, threats to soil like organic matter decline, erosion or compaction processes, water stocks management should be addressed urgently for preserving future agricultural resources.

The purpose of the DIGISOIL project was to develop pertinent, reliable, and cost-effective geophysical solutions for mapping soil properties. Considering recent advances in equipment and signal processing, collecting soil information at the catchment scale needs the deployment of geophysical sensors of different nature. The identified sensing systems included the ground-penetrating radar (GPR), electrical resistivity tomography (ERT), electromagnetic induction (EMI), seismics, and airborne hyperspectral). Some tests were also realized with magnetics. This panel of methods are at the basis of the DIGISOIL's mapping tools designed to provide inputs to DSM applications. The activities presented here were thus focused on (i) adapting, testing, and validating the geophysical technologies, (ii) establishing correlations between measured geophysical parameters and soil properties involved in soil functions/threats such as erosion, compaction, organic matter decline, (iii) evaluating the societal impact of the developed tools by investigating their relevance to end-user needs, their technical feasibility, and their cost-effectiveness.

Contents

List of illustrations

List of tables

1. Introduction

1.1. CONTEXT AND PROBLEMATICS

Since 2002, the European Commission considers the soil as a non-renewable resource threatened by natural hazards and human activities. Erosion, organic matter decline, contamination, sealing, compaction, salinisation, floods and landslides have been identified as the main soil threats. The implementation of the Soil Thematic Strategy (COM232, 2006) implies that Member States individuate risk areas for erosion, organic matter decline, compaction, salinisation and landslides. For soil contamination, an inventory of contaminated sites should be undertaken. Concerning risk area delineation (Eckelman et al., 2006), soil and soil-related parameters – referring to as auxiliary data by the digital soil mapping (DSM) community – should be considered in the mapping model. These parameters also play an important role in the determination of many of the soil functions as well as in the characterisation of the current status of soil degradation. Among the main soil and soil-related parameters to be considered by Member States for delineating risk areas (COM232, 2006), soil texture, soil water content – and related hydraulic properties, bulk density and soil organic matter, impact crucially many soil functions, so that these properties have to be considered as a first priority. In particular, threats to soil like organic matter decline, erosion or compaction processes, water stocks management should be addressed urgently for preserving future agricultural resources (Tóth, 2008).

1.2. OBJECTIVES

The purpose of our project was to develop pertinent, reliable, and cost-effective geophysical solutions for mapping soil properties. Considering recent advances in equipment and signal processing, collecting soil information at the catchment scale needs the deployment of geophysical sensors of different nature. The identified sensing systems included the ground-penetrating radar (GPR), electrical resistivity tomography (ERT), electromagnetic induction (EMI), seismics, and airborne hyper-spectral). Some tests were also realized with magnetics. This panel of methods are at the basis of the DIGISOIL's mapping tools designed to provide inputs to DSM applications. Our objectives were thus focused on (i) adapting, testing, and validating the geophysical technologies, (ii) establishing correlations between measured geophysical parameters and soil properties involved in soil functions/threats such as erosion, compaction, organic matter decline, (iii) evaluating the societal impact of the developed tools by investigating their relevance to end-user needs, their technical feasibility, and their cost-effectiveness.

1.3. STATE OF THE ART

1.3.1. Geoelectric

Widely developed for the characterisation of the unsaturated zone, measurements of the electrical resistivity have recently been applied to characterize soils (Banton et al., 1997) by recording the difference in electrical potential in the soil when a quasi-continuous current is injected. The electrical resistivity depends on several soil properties and has been used i) to characterise the texture of the soil, especially its clay content of soil (Giao et al., 2003), ii) to describe the pore geometry especially with the help of the Archie's law (Archie, 1942), iii) to measure the soil water content (Binley et al., 2002) and its temporal variation, and iv) to discuss the salinity of the soil or the composition of the soil solution (Kalinski et Kelly, 1993). In the field, 2D or 3D tomographies enable to describe the variation of soil porosity or soil density (Besson et al., 2004) and the evolution of the water content (Michot et al., 2003). The apparent resistivity data from these 2D or 3D electrical tomographies can then be interpreted by means of established softwares (Loke and Barker, 1996) whatever the expected resolution – from few centimetres to few meters. Recent technological improvements have consisted in building automatic profilers that can be used in the field for prospecting areas up to a few hectares: this technology is now operational and consists in a continuous multi-depths electrical resistivity profiler, called MUCEP (Panissod et al., 1997). The apparent electrical resistivity is measured for three pseudo-depths and its interpretation helps in delineating different soils in digital soil mapping, without any inversion of the signal (Tabbagh et al., 2000). Nevertheless several problems still limit the interpretation of electrical resistivity measurements in terms of soil properties. First, the inversion of apparent data can diverge when there are sharp variations of electrical resistivity values; this situation is really common when the soil dries and when cracks develop at the soil surface, even for clay content below 20 %. Second, the electrical resistivity signal can be influenced by several soil properties simultaneously, which limits its interpretation in one property, for example the effect of bulk density and water content. Finally, the resolution of the interpreted signal decreases versus depth and is usually not precise enough to describe the soil horizons. Other geophysical techniques have then to be taken into account for that interpretation or for joined inversion.

1.3.2. Seismics

Due to the development of subsurface characterization studied for environmental purposes, the efficiency of seismic methods for estimating ground velocity structures and mechanical properties has been in real progress and has found various applications in the field of waste disposal (Lanz et al., 1998), landslides (Grandjean et al., 2007), or hydrogeophysics (Sturtevant et al., 2004). New equipment with 48 or 72 channels and PC-piloted acquisition software contributed to the development of this method, for example by reducing the acquisition times with unplugged gambled sensors (Grandjean, 2006a). This improvement was also supported by the development of new data processing protocols like acoustical tomography (Grandjean, 2006b) or Spectral Analysis of Surface Waves (Park et al., 2000; Grandjean and Bitri, 2006). The adaptation of seismic methods to soil properties mapping is conditioned by

the possibility to reduce the seismic antenna (originally of several tens of meters) to around several meters, in order to investigate the first two meters of the ground. This implies to work with very high source frequencies for dealing with high resolution data.

1.3.3. GPR

Ground-penetrating radar (GPR) is an increasingly used non-invasive and proximal electromagnetic (EM) sensing technology to image the subsurface and identify its physical properties (Daniels, 2004). It is based on the radiation of ultrawide band VHF-UHF EM waves into the soil and recording of the reflected signals. In the areas of agricultural and environmental engineering, GPR has been used to identify soil vertical structure (Grandjean et al., 2001), to locate water tables (Nakashima et al., 2001), to follow wetting front movement or soil hydraulic parameters (Cassiani and Binley, 2005). Yet, existing GPR techniques still suffer from major limitations due to the strongly simplifying assumptions on which they rely with respect to EM wave propagation phenomena. To circumvent these shortcomings, Lambot et al. (2004) have recently developed a new approach based on stepped-frequency continuous-wave monostatic off-ground GPR. With such system, based on ultrawide band vector network analyzer (VNA) technology, the measured physical quantity is exactly known and is defined as an international standard. This permits the use of advanced full-wave forward and inverse modelling techniques to estimate the soil EM properties (Lambot et al. (2004b) This approach has been successfully validated in a series of controlled hydrogeophysical experiments for soil EM characterization, including dielectric permittivity, electric conductivity and frequency dependence of these quantities (Lambot et al., 2006).

1.3.4. Hyperspectral

Visible and Near InfraRed (VNIR) Spectral analysis and diffuse reflectance analysis are techniques to rapidly quantify various soil characteristics simultaneously (Malley et al., 2004). Numerous studies used laboratory spectrometers to analyse SOC (McCarty et al., 2002) and show notably their relevance for SOC inventories. Further, using the NIRS concept, Ben-Dor et al. (1997) were able to show that degradation stage of soil organic matter (SOM) can be assessed solely from lab spectroscopy. On the other hand, imaging spectroscopy, or hyperspectral remote sensing, has been mainly used for the mapping of surface soil properties of agricultural fields with high spectral and spatial resolutions (Ben-Dor et al., 2002; Stevens et al., 2006). It is important to note that SOM is a property consisting of significant chromophores across the entire spectral region and hence is easy to detect via NIRS approach at the laboratory level. Most authors, however, have not investigated or considered the problems encountered during applying NIRS under imaging condition (e.g. sampling representation, scene dependency or atmospheric attenuations) and hence, questions on how each level is influenced by the measurements conditions, remain open. Previous studies, conducted within the PRODEX program (sensors: CASI-SWIR 2002, CASI-ATM 2003, AHS 2005), have explored the capabilities of VNIR spectroscopy, in the context of SOC inventories and monitoring. It has been shown that the pre-treatment of spectra with appropriate mathematical algorithms, e.g. Savitzky-Golay, and the use of Partial Least

Square Regressions (PLSR) are efficient techniques to calculate the carbon content from the spectra, but the problem of disturbing factor, like soil moisture and roughness, need nevertheless be addressed.

1.4. WORK PROGRAMME

In the following, we present the outcomes resulting from several activities carried out to fulfil above-cited objectives. These activities are related to studying pertinent sensors and measuring technologies, to estimating the capabilities of geophysical techniques and processing algorithms for investigating highly complex soil properties and to capability to draw up soil properties maps. A series of field validations, realised on two test-sites are presented, taking into account the uncertainties issues. Finally, the estimated soil properties maps are interpreted in terms of threats to soil or soil functions.

2. Instrumental developments and methods

2.1. RESISTIVITY TECHNIQUES

Resistivity techniques were operated by the mean of the MUCEP sensor (*Figure 1*). This device was initially developed by CNRS Garchy (Panissod et al., 1997). The system is composed by a SYSCAL acquisition system and a series of electrodes mounted on metallic wheels; a Doppler radar triggers measurements every 10 cm along parallel profiles. The system can be towed by a tractor or a quad, the positioning of the data points being done by a differential GPS system. The sensor part shows 3 quadrupoles configured in a “V shape” mode for optimizing signal to noise ratio. Electrodes spacing are 0.5, 1 and 2 m for each electrode quadrupoles respectively. While the system is moving, measurement are realised in a continuous mode by the mean of a 150 milli-second time pulse. In the same time, signals are recorded according to the 3 acquisition channels of the SYSCAL station. This equipment allows 1,000V (2,000V peak to peak). In all cases maximum current is 2.5A, with maximum power of 250 Watt.

Figure 1: The MUCEP system: The data acquisition station is located inside the vehicle; The first two wheels carry the injection electrodes while 3 the last pairs carry measurements channels for 3 different spacings

Data are interpreted according to a very simple earth model (1D) for each spatial position where measurements are performed. The inversion can be done by using the QWIN1D software that uses the Levenberg-Marquadt optimisation algorithm (Cousin et al., 2009). The most important soil properties the electrical resistivity can provide are

soil water content and soil texture. For both cases, the principle of retrieving these soil properties consists in using correlations that link these values. Such correlation can be obtained from soil samples as shown in Figure 2:

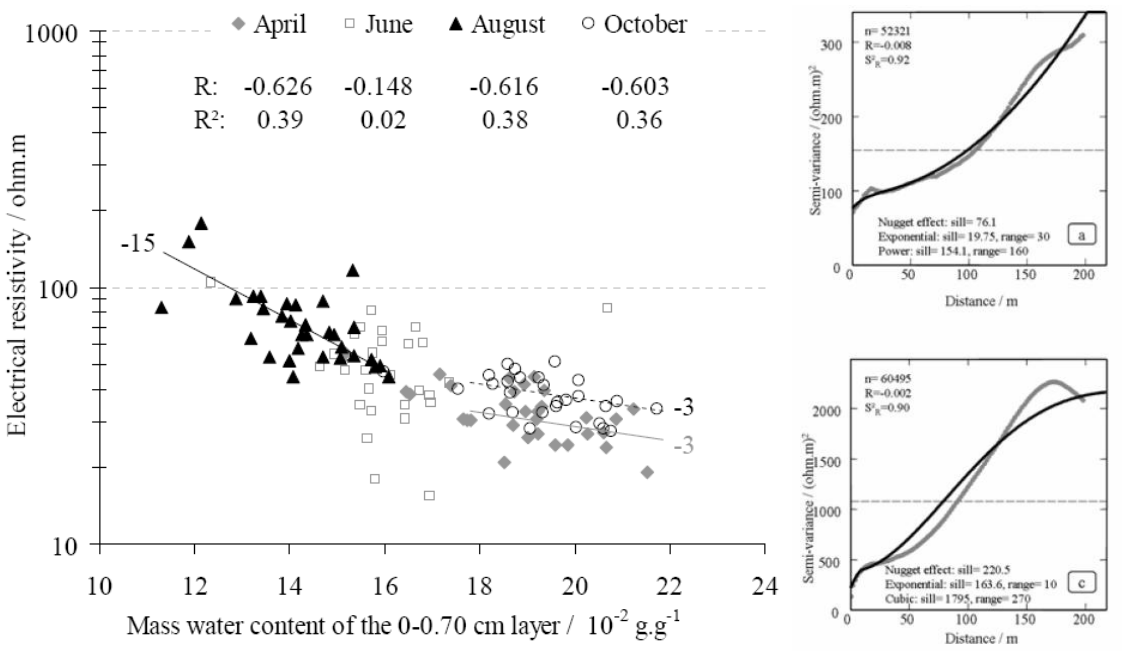


Figure 2 : LEFT: Relationship between the electrical data from the MUCEP Channel 1 and the water content measured in the 0-0.70 m soil layer. RIGHT: Experimental (grey) and theoretical (black) variograms of the electrical resistivity measured at MUCEP channel 1 with the cross-validation indicator at two seasons. From Besson (2007).

2.2. GROUND-PENETRATING RADAR

In DIGISOIL, a frequency-domain, zero-offset, off-ground GPR prototype set up with vector network analyzer (VNA) technology and an off-ground directive horn antenna was particularly tested (Figure 3). A major advantage of the VNA over traditional systems is the accurate calibration of the system, which is an international standard, thereby ensuring proper repeatability of the measurements. In addition, VNA technology permits to have a full control on the GPR frequencies that are transmitted (stepped-frequency continuous-wave (SFCW) system). The off-ground mode is used here as it permits accurate full-waveform forward and inverse modelling of the radar data for quantitative reconstruction of the soil properties.

Figure 3: VNA ZVL from Rohde & Schwarz → SFCW radar, double ridged horn antenna BBHA 9120F from Schwarzbeck (Antenna dimensions : 96 x 68 x 95 cm and frequency range : 200-2000 MHz) The mode of operation is zero-offset, at 1.1 m above the ground.

The common surface reflection method applies to off-ground GPR configurations, either monostatic or bistatic, and is based on the determination of the Fresnel reflection coefficient of the soil surface interface. The following assumptions are particularly considered: (1) the antennas are located in free space (air) above a homogeneous half-space (soil) limited by a plane interface, (2) the reflection coefficient can be approximated by the plane wave reflection coefficient, (3) antenna distortion effects are negligible, (4) the soil electrical conductivity is assumed to be negligible, (5) the magnetic permeability is assumed to be equal to the free space permeability, and (6) the dielectric permittivity is frequency-independent. As a result, the reflection coefficient at the soil interface is a Dirac's delta function of time and its amplitude is defined as the ratio between the backscattered (E_s) and incident (E_i) electric fields. For a normal incidence plane wave, the amplitude R of the reflection coefficient can thus be expressed as:

$$R = \frac{1-\sqrt{\varepsilon_r}}{1+\sqrt{\varepsilon_r}} \tag{1}$$

where ε_r is the relative dielectric permittivity of the soil. The soil dielectric permittivity can therefore be derived as:

$$\varepsilon_r = \left(\frac{1-R}{1+R}\right)^2 \tag{2}$$

The reflection coefficient R is usually determined from the measured amplitude of the soil surface reflection, A, relative to the amplitude measured for a perfect electric conductor (PEC) situated at the same distance as the soil, namely, A_{PEC}. The ratio between the reflection coefficient at the soil surface interface (R) and at a PEC interface (R_{PEC}) can be expressed as:

$$\frac{R}{R_{PEC}} = \frac{\frac{E_s}{E_i}}{\frac{E_{s,PEC}}{E_i}} \tag{3}$$

Since R_{PEC} = -1, assuming E_i to be constant, and assuming that the measured amplitude A is directly proportional to the backscattered electric field E_s, i.e., there are no antenna distortion effects, Equation (3) reduces to:

$$R = -\frac{E_s}{E_{s,PEC}} = -\frac{A}{A_{PEC}} \tag{4}$$

Notwithstanding the practical appropriateness of this method for mapping applications, the method has been applied by a few authors only. A major limitation of the method is related to the requirement of PEC measurements with the antenna exactly at the same height as for the measurements above the soil, without which significant errors are introduced. The concept is however commonly used in airborne and space-borne radar remote sensing for the retrieval of soil surface water content. When the soil dielectric permittivity is known, it is correlated to the volumetric water content using a petrophysical model such as Topp's equation (Figure 4).

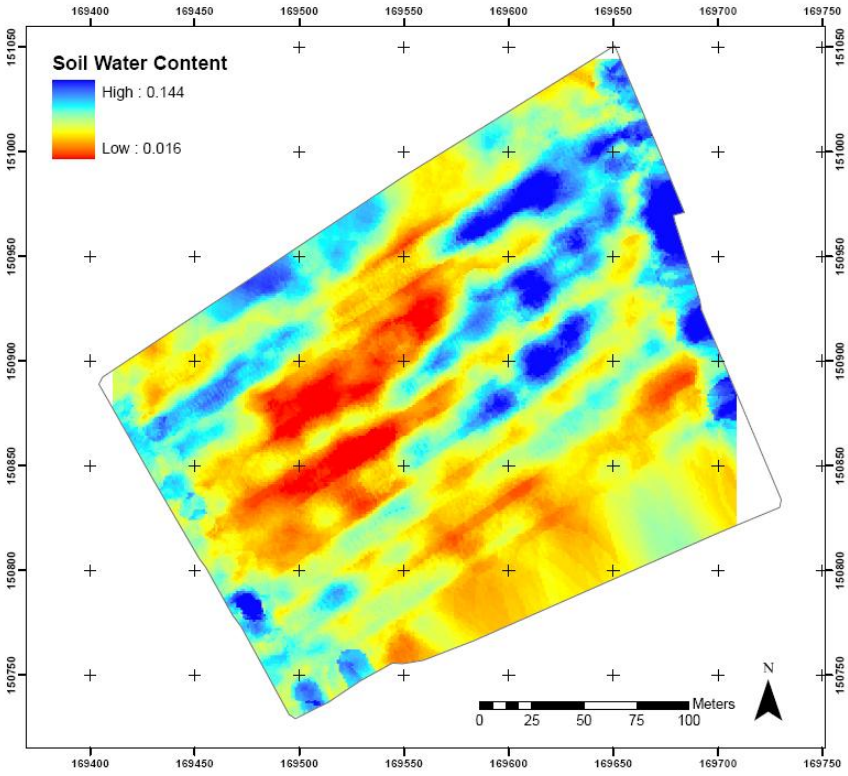


Figure 4: Soil surface water content estimated from off-ground GPR data using full-waveform inverse modeling (UCL test site).

2.3. SEISMIC TECHNIQUES

Seismic sensors used in DIGISOIL are tractable gambles of 10 Hz central frequency from Mark Products. Optimal sensor spacing that allows a precise velocity picking on the dispersion diagram has been estimated to 50 cm because the maximum energy can be clearly observed on the whole bandwidth and propagation modes can be distinguished with accuracy. The total seismic acquisition system is composed by a hammer striking on an anvil for generating the source signal, a seismic antenna of 24 10 Hz gambles spaced of 50 cm, a Geometrics GEODE unit for the numerical management of the diverse signals and a PC-based central unit for piloting the acquisition using the Geometrics recording software (Figure 5).

Figure 5 : Seismic device mounted on caterpillar-type vehicle for mapping of the soil shear waves velocity. The sensor system is mounted at the back of the vehicle. A computer automatically controls the measurements.

The spectral analysis of surface waves (SASW) method is an *in situ* seismic, non destructive technique used for evaluation of layers thickness and the associated shear waves velocity (Vs) with depth in layered systems. We focus on the exploitation of surface waves by analysing the dispersion behaviour of these waves, Three steps are involved in a surface wave test (Figure 6): (1) field testing for recording surface waves, (2) determination of the experimental dispersion curve from the field data, and (3) inversion of shear wave velocity profile from the experimental dispersion curve. A 2D wavefield transform method is used to determine experimental dispersion curve. It transforms data from space–time domain into a more convenient domain where the dispersion curve, i.e., the plot of phase velocity versus wavelength (or frequency), can easily be identified McMechan and Yedlin (1981). This method first performs a p-t stack followed by a transformation into the p-ω domain. The Fourier spectrum of a seismic signal at a distance r_n being

$$A(\varpi, r_n)e^{i\phi(\varpi)_n}$$

One possible p-ω stack of N traces at different distances from the same source is defined by the relation

$$F(p,\varpi) = \sum_{n=1}^{N} C(\varpi, r_1, r_n)^{-1} A(\varpi, r_n) e^{i\phi_n} e^{i\varpi p r_n}$$

Where

$$C(\varpi) = A(\varpi, r_1, r_n)e^{i\phi_1}\sqrt{\frac{r_n}{r_1}}$$

Searching for the maxima of |F|, yields the possible dispersion curves. These curves are afterward inverted to obtain the related Vs vertical profile. The inversion process consists in finding a set of parameters x which minimizes these residues in the least square sense. The parameters x are linked to the data y through a model, which consists in a nonlinear relation (Tarantola, on 1987). This nonlinear problem can however be linearized to an iterative problem defined by:

$$(Ax - y)^t (Ax - y) + \sigma^2 x^t x = MIN$$

Where A is the matrix of the partial derivatives of phase velocity with regard to the parameters (Vs and thicknesses of layers) and where σ^2 contains the a priori information about the model. The matrix A can be decomposed as a function of its values (Λ) and main vectors (U and V) in the following way:

$$A = U\Lambda V^t$$

Then, the solution of the problem by decomposition in singular values is:

$$x = V\left(\Lambda^2 + \sigma^2 I\right)^{-1} \Lambda U^t y$$

The inversion process is implemented according to Herrmann (1987).

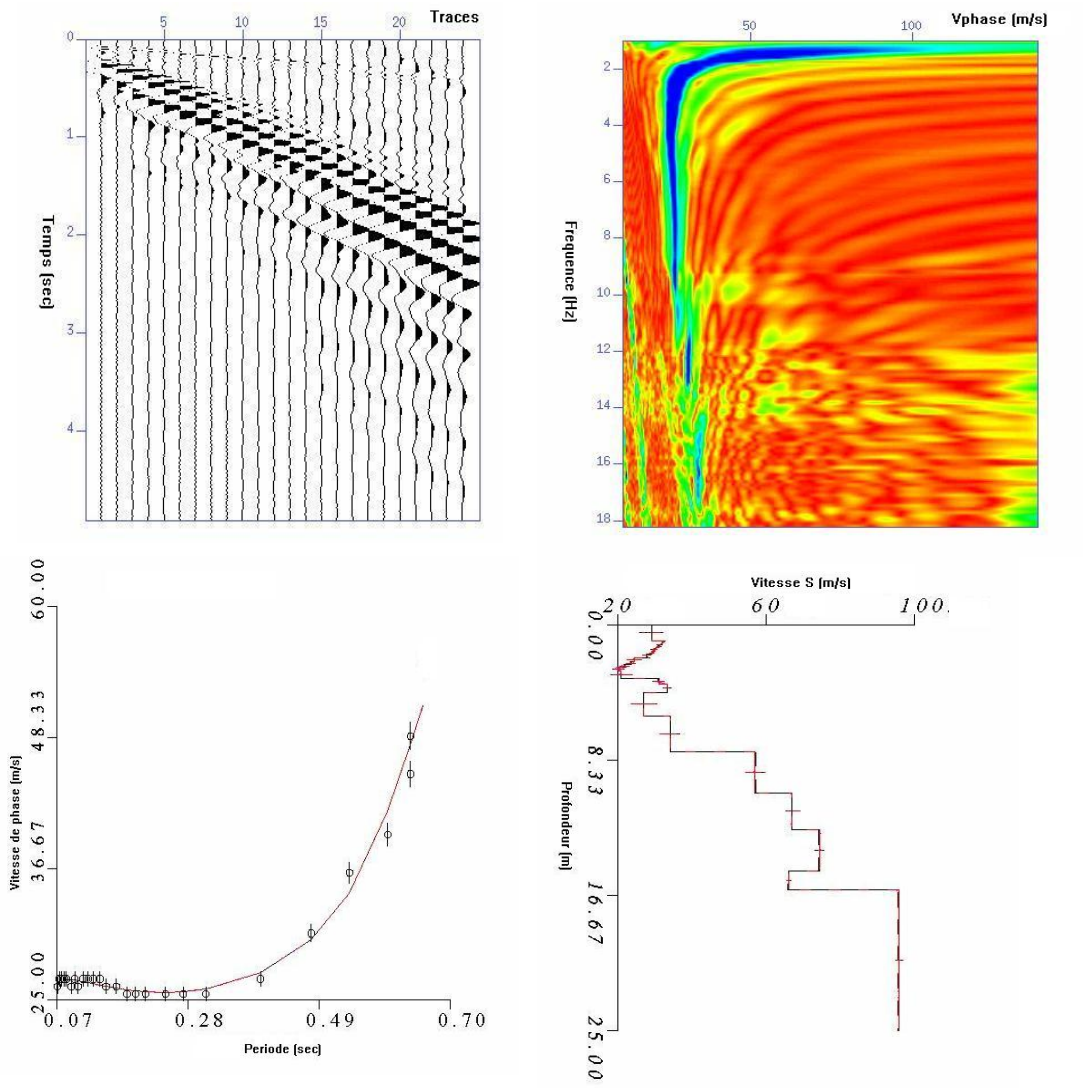


Figure 6: a-Example of short record with obvious low velocity Rayleigh wave (b)- Phase velocity plot obtained by interception time ray parameter transform c- Experimental and theoretical phase velocity fitting by linear inversion d-Depth variation of shear velocity

2.4. HYPERSPECTRAL

The Galileo Avionica SIM.GA HYPER is a 512 + 256-spectral-band push-broom sensor with VNIR and SWIR imaging capability. The airborne hyperspectral system covers the 400-2450 nm spectral region and is operated for DIGISOIL campaigns on board of the UNIFI ultra-light aircraft FOLDER at about 1000 m of altitude. The hyperspectral HYPER SIM.GA is composed of two optical heads (Figure 7):

1) VNIR Spectrometer with a spectral range of 400-1000 nm, 512 spectral bands with 1.2 nm spectral sampling, 1024 spatial pixels across a swath of 722 m (@ H= 1000m), which corresponds to a pixel resolution of 0.7 × 0.7 m

2) SWIR Spectrometer with a spectral range of 1000-2450 nm, 256 spectral bands with 5.8 nm spectral sampling, 320 spatial pixels across a swath of 425 m (@ H= 1000m), which corresponds to a pixel resolution of 1.33 × 1.33 m

The optical heads are managed by a common data acquisition and control electronics. The HYPER SIM.GA works as a push-broom imager. A spatial line is acquired at nadir and the image is made exploiting the aircraft movement. The optical head of HYPER SIM.GA is rigidly coupled to a GPS/INS unit that collects data about platform movements (yaw, roll, pitch, velocity, altitude, lat, long) allowing to geo-rectify the images acquired. The use of GPS/INS unit reduces the mass and the cost of the instrument avoiding stabilized platform. Given the aircraft velocity, it is possible to get the minimum altitude required to acquire the scene without under-sampling. The relationship is: H/v >25s, where: H is the aircraft altitude in meters and v is the velocity in m/s. The actual instrument configuration allow to state that HYPER is best suited for studies/campaigns in which high spatial and spectral resolution are priorities (spatial/spectral discrimination as well as simulation of new medium resolution products) compared to the wide viewing for intensive mapping needs.

Figure 7: HYPER SIM.GA and UNIFI FOLDER installation and bottom view of optics

In DIGISOIL, all imagery were processed according to the Galileo Avionica processing chain generating and distributing the defined radiance and reflectance data set over the study areas. Briefly the main generated products were as follows:

- Raw data (L0a) include separate VNIR and SWIR raw DN data blocks are in ENVI BIL format; navigation INS/GPS data are splitted and time- frame synchronised for each image block; in-flight dark data acquired before and after each flight line.
- Georeferenceable at-sensor radiance (L1a): VNIR/SWIR DN values are converted to at-sensor radiance applying key data parameters and calibration coefficients from laboratory measutrements. File are still ENVI BIL format but scale factor of 1.0e-4 is applied to fit an unsigned integer data type, output units are W/(m2 sr nm).
- Geocoded products were derived by using the PARametric Geocoding PARGE software integrating for each flight line the GPS position, and roll-pitch-heading

attitude from the CMIGITS-III unit and the resampled DEM for the HYPER pixel. The PARGE outputs for each HYPER image are:

*igm file (Lat/Long Geographic LUT for ENVI)

*sca file (scan zenith/azimuth angles and altitude for ENVI)

There are two approaches generally adopted to retrieve a given property from remotely acquired spectral data. The physical approach involves the inversion of a Radiative Transfer Model (RTM) describing accurately the reflectance of an object as a function of its surface geometry and chemical/physical characteristics. Such models, based on physical laws, establish a direct link between a given property and spectral variations of an object and therefore offer the possibility to quantify its properties in a robust way. While RTM have been widely applied in the remote sensing literature to retrieve vegetation characteristics (see e.g Jacquemoud et al.,2000; Atzberger, 2004; Darvishzadeh al., 2008), similar studies are less common for soil applications (see, however, e.g. Cierniewski and Karnieli, 2002). In the empirical/statistical approach, a correlation is established between the property of interest and the spectral information using calibration samples analysed with conventional methods. A so-called calibration model is then used to derive the values of unknown locations using their reflectance. A (pseudo-) independent test set is often collected to validate the model and assess its predictive ability. Soil reflectance is the result of complex interactions between incoming solar radiations, surface characteristics (roughness, vegetation residues) and soil physical/chemical properties (SOC, moisture, clay, iron oxides, etc…). Therefore, a direct interpretation of a soil spectrum is hard to achieve and the quantitative prediction of soil properties must resort to complex statistical models. The specific nature of spectral data (high-dimensionality and multi-collinearity) requires indeed the implementation of specific multivariate calibration tools. Numerous multivariate techniques such as Multiple Linear Regression (MLR, e.g. Dalal and Henry, 1982; Ben-Dor et al., 2002), Artificial Neural Networks (ANN, e.g. Fidêncio et al., 2002a; Daniel et al., 2003), Multivariate Adaptive Regression Spline (MARS, e.g. Shepherd and Walsh, 2002), Regression Trees (RT, e.g. Cohen et al., 2005) Principal Component Regressions (PCR, e.g. Chang et al., 2001; Islam et al., 2003) Partial Least Square Regressions (PLSR, e.g. Fidêncio et al. 2002b, Reeves and Delwiche, 2003) have been used so far to relate spectral measurements to soil properties. We will focus here on PLSR which was developed by Wold et al. (1983) and is one of the most common algorithms used in the literature. The algorithm is available in several commercial and non-commercial statistical softwares (e.g. Unscrambler, SAS, R, Matlab). A description of the PLS algorithm is given by Geladi and Kowalski (1986). The general PLS model has the form:

$$X = TP' + E$$
$$Y = UQ' + F$$

with Y, the matrix of responses, X, the matrix of predictors, T, X-scores, P, X-loadings, E, X-residuals, U, Y-scores, Q, Y-loadings, and F, Y-residuals. Similarly to PCR, PLSR involves the extraction of latent variables, i.e. linear combinations of initial explaining

variables. However, in PCR no information on Y is exploited to calculate these latent variables. Therefore, there is no guarantee that they are relevant to explain variation in Y. The PLS approach seeks linear combinations of the predictors that maximize the covariance between X-scores and Y-scores and therefore that explain both response and predictor variation. The optimal number of latent variable in a model is usually determined by minimizing the value of the Predicted Residual Sum of Squares (PRESS) based on *leave-one-out* cross-validation.

Multivariate calibration methods such as PLSR have been typically applied to spectral data acquired in the laboratory and under controlled/stable measuring conditions. When spectral data is acquired in the field or with a remote sensor, environmental conditions cannot be controlled anymore. In comparison with physically-based approaches, statistical methods are therefore considered as site- and sensor-specific due to spatio-temporal variation in soil characteristics not related to the studied property such as roughness, moisture content, clay types, etc…. (e.g. Stevens et al., 2008). As a result, it is often difficult to develop a universal calibration model. An alternative to such complex statistical approach is to use simple spectral indicesfor example to retrieve soil moisture (Whiting et al., 2004) and SOC content (Bartholomeus et al., 2008). Two examples are given in Figure 8 showing a stable relationship between laboratory- or field-based spectral indices and the target variable ($CaCO_3$ and moisture content). Such indices are therefore likely to show good calibration transferability between sensors or different measuring conditions.

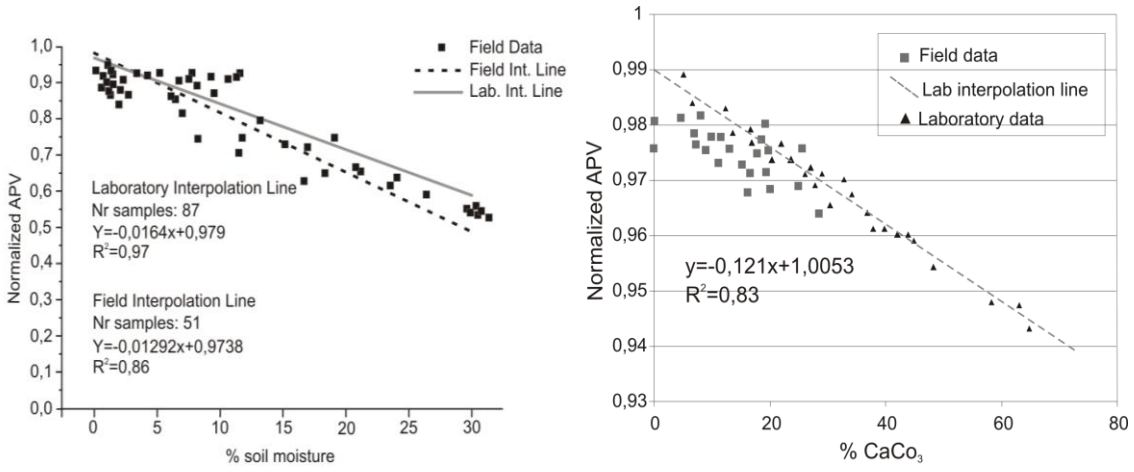


Figure 8: (a) Normalized absorption peak values (at 1460 nm) vs soil moisture (%). (b) Normalized absorption peak values (at 2436 nm) vs $CaCO_3$ content (%). In the laboratory, weight percentages of 5, 10, 15, 20 and 25 % of water were added to dried samples. Soil samples were also mixed with increasing amounts of a standard calcite powder (purity higher than 98%). The total concentration of $CaCO_3$ was calculated by adding the natural $CaCO_3$ content in the samples, determined by calcimetry.

2.5. DATA INTEGRATION

In the context of the DIGISOIL project, five soil characteristics have been identified as parameters of interest, for developing indicators dealing with i) compaction, ii) decrease

in organic matter, iii) erosion iv) shallow landslides. These soil characteristics have already been identified as: bulk density, texture (clay content), carbon content, water content and an additional one related to horizonation describing the fact that all the former parameters vary with depth along a soil profile (Figure 9).

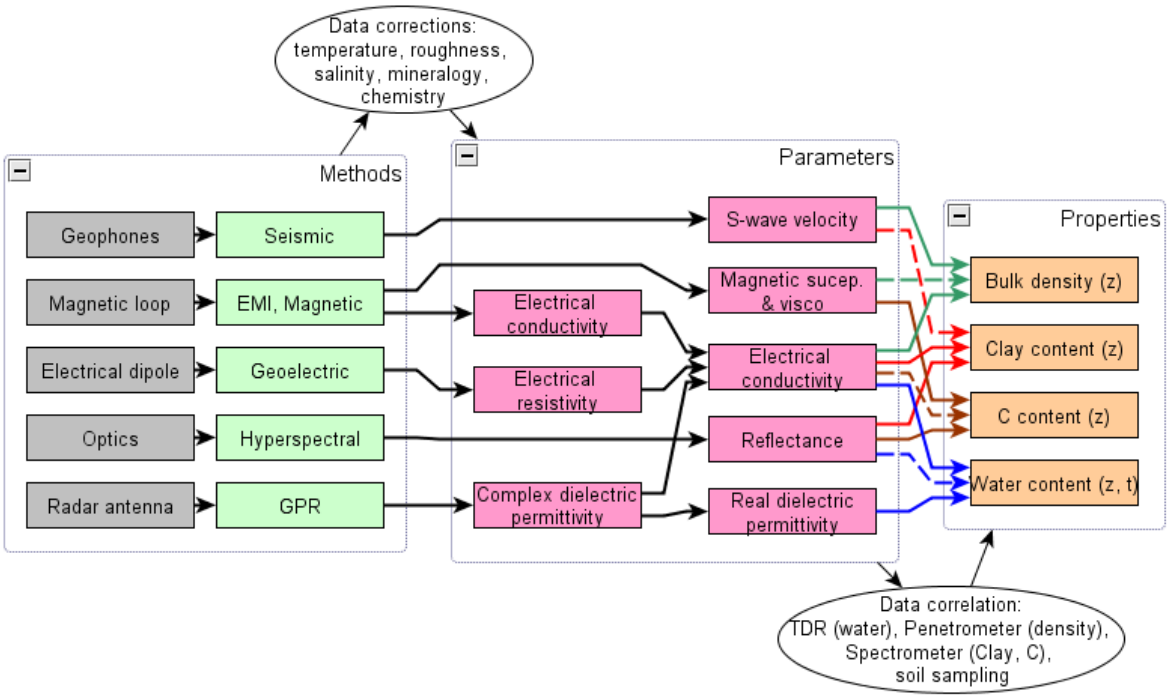


Figure 9 : Diagram showing the different paths for going from sensors to soil properties.

The transformation processes suggested by the Figure 9 are most of the time not straightforward and need to tackle some issues that will be discussed in the next chapters when processing flows carried out on each dataset will be described. The two main processes are the followling

- Correction to be applied: some of the geophysical parameters, inverted from field measurements, are strongly dependent on time physical variables like temperature, water content, etc. When applicable, these corrections will be described and discussed in the framework of the next field measurement activities;

- Data calibration and assimilation: when transformations are more complex than simple physical or empirical relationships, more sophisticated methods, based on correlation or fusion strategies, can be used. These methods will be listed afterwards and described according to each of the geophysical parameters studied;

3. Experimentations

3.1. THE LUXEMBOURG SITE

3.1.1. Site presentation

The pilot zone (Figure 10) consists of a number of cropland fields (c. 30) along a transect in the Grand Duchy of Luxembourg to cover a wide range of parent materials, soil types, SOC content and ranging from extensive livestock farms to arable farms with low manure input. Although of a small extent (2586 km2), Luxembourg is characterized by a very diverse physiogeography. Five distinct agrogeopedological zones can be distinguished: (1) the Oesling in the north covering one third of the country is a homogeneous schist Hercynian massif with a mean altitude of 450 m. Main soil texture is sand and loam. The southern part of Luxembourg, called the Gutland, covers two thirds of the country. In this area, four different agro-geopedological zones are distinguished: (2) the Minette basin in the south is characterized by red loamy/clay soils, formed by marls and sandstones covered with ferrous sediments; (3) the Moselle region characterized by Keuper and limestone with clay and lime soil types; (4) central part of the Gutland in the Luxembourgish sandstone area with sandy soils; (5) the Rédange-Diekirch area with soils of the loam-loess type on red sandstone and Luxembourgish sandstone. In the southern part of Luxembourg (Gutland), the most common crop rotation is a three years rotation with winter wheat, winter barley and silage maize. The field on which all geophysical experiments were carried out is in the Northern part of Luxembourg (Oesling). The field of 5 ha is slightly sloping to the West starting on a plateau. Soils are generally thin and stony.

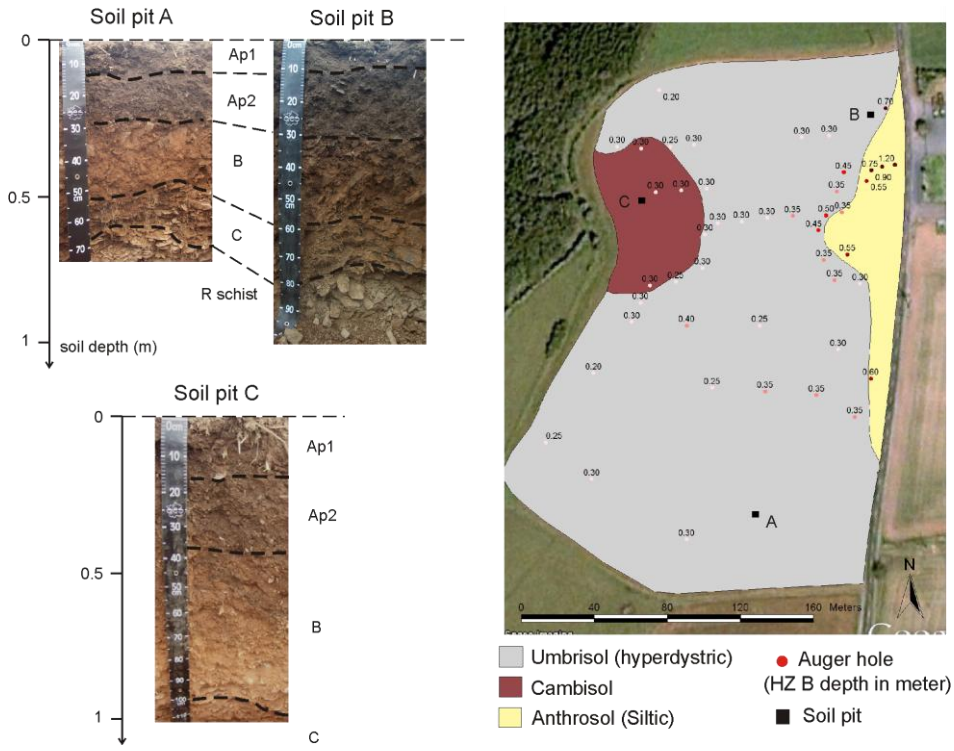


Figure 10 : Pedologic desciption of the studied area. The value associated with dot is the depth of the B horizon.

3.1.2. Experimental setup

Five types of measurements and data were collected on field: geoelectric, seismic, hyperspectral, electromagnetic induction (EMI) and off and on-Ground Penetrating Radar (GPR) data (Figure 11 and Figure 12). For the Luxembourg site, it means that:

- Electrical resistivity measurements were obtained at the field scale by the use of a device similar to the MUCEP device described in the section 2.1 of this report and called ARP (Automatic Resistivity Profiling, Geocarta society). ARP system consists in simultaneous measurements of electrical resistivity for 3 investigation depths. The difference with MUCEP is the distance between the current injection electrodes and the resistivity measurements electrodes of the V3 array: 1.7 m instead of 2.0 m. In addition, 5 vertical electrical soundings (VES) were realized at different locations in the plot;

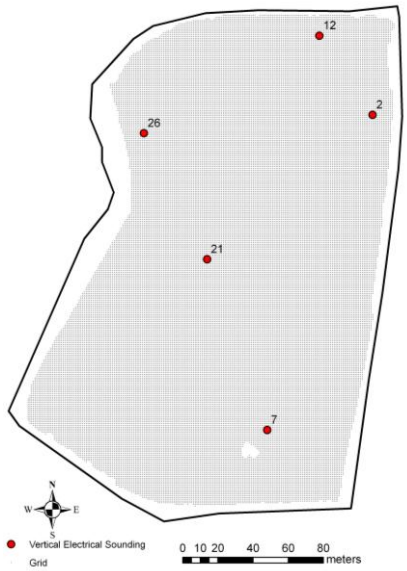


Figure 11 : Location map of the electrical resistivity measurements: grid of ARP measurements and locations of VES

- Seismic experiments led to realize 5 profiles covering around 5 ha for almost 200 seismic shots. In addition, 30 penetrometers were performed on the same area to validate the method;

- Hyperspectral data were acquired from the AHS 160 airborne sensor in order to produce the reflectance signal of the bare topsoil. This signal was then correlated with the C content of the plough layer (0-20 cm). 27 topsoils (0-20 cm) have been sampled in the selected and an adjacent field;

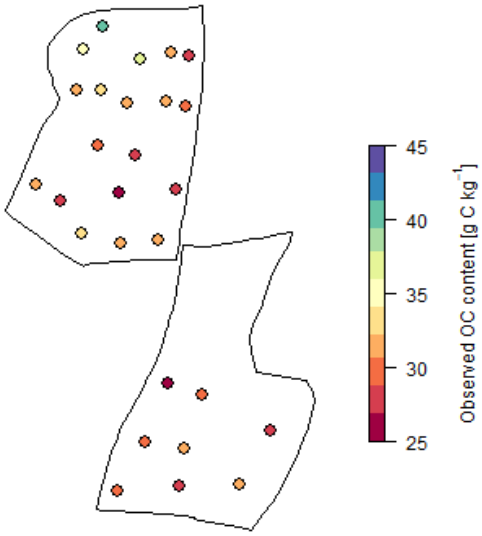


Figure 12 : Calibration points and their C content in the topsoil (0-20 cm)

- Electromagnetic induction (EMI) data were acquired with two sensors: the EMP-400 Profiler (GSSI) and the EM38 (Geonics). The Profiler allowed to perform measurements at three frequencies (5, 10 and 15 kHz) simultaneously while the EM38 was mounted on a quad used as a platform for both GPR and EMI measurements. EMI measurements were carried out along transects spaced approximately 15 m apart, with one transect on two passing close to the locations of the ground truth measurements. The resolution (measurement spacing) within transects was about 1 m. Several sets of measurements were performed over two days;

- The combination of a Vector Network Analyzer (VNA) and a horn antenna (200-2000 MHz) located at 1.1 m height, was used to perform the off-ground measurements. The system was mounted on a quad and a differential GPS was used for data geo-referencing. The data acquisition was controlled by a portable computer which recorded measurements every 1 second. The measurements were repeated during two days over the entire study area and performed along S-N and N-S transects at the same locations as the EM38 measurements;

- A time-domain GPR system (model SIR-20, Geophysical Survey System, Inc., GSSI, Salem, Massachusetts, USA) combined with a pair of 400 MHz shielded bow-tie antennas located at 1.1 m distance from each other was used as impulse radar. Applying the two antennas allowed the use of both monostatic and bistatic modes simultaneously. On-ground GPR measurements were performed following the same transects as for the off-ground radar measurements;

- The magnetic properties measurements were located on 231 datapoints. The magnetic susceptibility was measured with 2 devices in 4 configurations (Bartington MS2D, CS60 VCP, CS60 HCP and CS60 VVCP). The magnetic viscosity was measured with 3 devices (ELSIEC DECCO, Protovale TS6 and VC100).

3.2. THE MUGELLO SITE

3.2.1. Site presentation

In Central Tuscany, and more particularly, in the hillsides north of Florence, soils with agricultural suitability have a high economic value mainly connected with the production of internationally famous wines and olive oils. Sediment yield and consequently soil losses are produced by two different mechanisms, erosion and landsliding, which affect all the types of land use. Even small disturbances induce potential economic losses which must be considered in farming management practices. The test area (Figure 13) is formed by a number of fields located in the Mugello basin, located about 30 km north of Firenze and it is extended for about 20 km^2.The geological terrains outcropping in these zones are fluvio-lacustrine deposits, ranging in age between lower Pliocene and Upper Pleistocene, forming low-dipping lenticular

beds. From the pedological point of view, more or less eroded soils prevail (utric and calcareous Regosols and lytic Leptosols), soils with pedogenetic structure at depth and weakly differentiated profiles (eutric and calcareous Cambisols), soils with clay masses (gleich Luvisols), acidic soils with accumulation of organic material (humic Umbrisols) and anthropically terraced soils (anthropic Regosols). Agriculturally suitable terrains are assigned mainly to annual crops, marginally to olive groves, vineyards and orchards. Climate is temperate-hot, with cold winters and mildly hot to hot summers and medium to intense rainfalls, concentrated in the autumnal months.

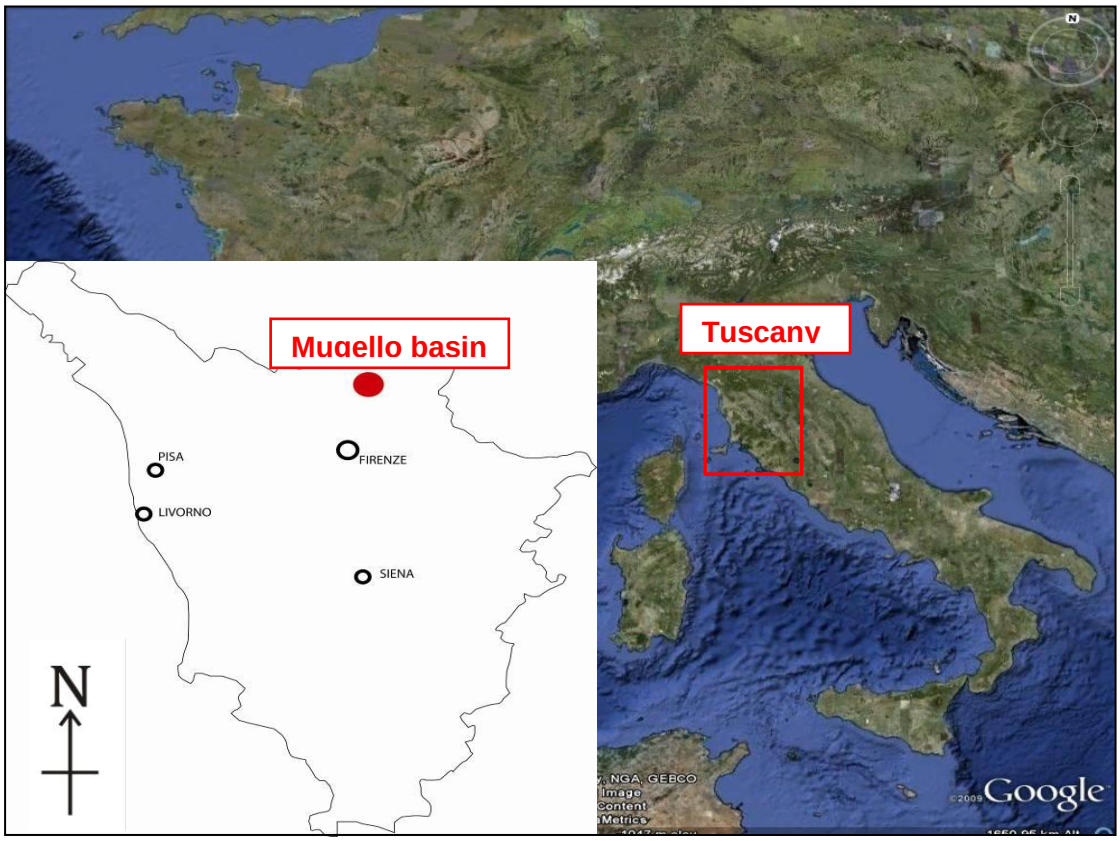


Figure 13 : Location of study area (Mugello basin).

3.2.2. Experimental setup

Five types of measurements and data were collected on field: geoelectric, seismic, hyperspectral, electromagnetic induction (EMI) and off and on-Ground Penetrating Radar (GPR) data. For the Mugello site, it means that:

- Electrical resistivity measurements were obtained at the field scale by the use of the ARP device. The measurements were realized along transects separated by 1 m covering the entire studied field. Along one transect, measurements

were performed and recorded every 10 cm. Local 2D and 3D Electrical resistivity Tomographies were realized;

- Seismic experiments led to realize 10 profiles covering few ha for almost 300 seismic shots. In addition, 10 penetrometers were performed on the same area to validate the method;

- A complete hyperspectral geocoded reflectance dataset was collected using SIM.GA hyperspectral image sensor from Selex-Galileo, mounted on board of the University of Florence ultra-light aircraft. The approximate pixel resolution was 0.6 m (VNIR) and 1.2 m (SWIR), considering a height of flight of about 900 m. Soil samples (to a maximum depth of 5 cm) were collected (positions recorded with a Leica 1200 differential GPS) simultaneously with the flight of SIM.GA hyperspectral camera and underwent traditional laboratory analyses (Rietveld XRD determination of clay content) and laboratory spectral signatures collection under controlled conditions, using an ASD FieldSpec spectroradiometer, provided with an illuminating device with stabilized krypton lamps, with fixed viewing and shooting geometry. The ASD spectroradiometer was used during flight as well, for ground measurements aimed at the characterization of atmospheric conditions (solar radiance and irradiance) and for reflectance measurements on reference targets such as metal panels and pcv sheets. A dedicated photogrammetric aerial survey was performed by UNIFI to obtain a 0.5 m precision digital elevation model, for orthorectification of SIM.GA images;

- The EMI setup was identical as that used in Luxembourg, namely, the EMP-400 Profiler (GSSI) instrument operating at three frequencies (5, 10 and 15 kHz), and the EM38 sensor mounted on a quad together with the GPR system. The data were geo-referenced using a differential GPS. As for Luxembourg, EMI measurements were carried out along transects spaced approximately 10-15 m apart, with about 1 m resolution within the transects. Data were acquired simultaneously with the EM38 in vertical dipole orientation and with the Profiler in horizontal dipole mode and vertical dipole Profiler measurements were performed the second day;

- As for Luxembourg, the combination of a Vector Network Analyzer (VNA) and a horn antenna (200-2000 MHz) located at 1.1 m height, was used to perform the off-ground measurements. The system was mounted on a quad and a differential GPS was used for data geo-referencing. The data acquisition was controlled by a portable computer which recorded measurements every 1 second. The off-ground GPR measurements were performed all over the studied area at the same locations as the EM38 measurements;

- As for Luxembourg, a time-domain GPR system (model SIR-20, Geophysical Survey System, Inc., GSSI, Salem, Massachusetts, USA) combined with a pair of 400 MHz shielded bow-tie antenna located at 1.1 m distance from each other was used as impulse radar. Applying the two antennas helped us to use the

monostatic and bistatic modes simultaneously. The on-ground GPR transects correspond to those followed for the Profiler in vertical dipole orientation mode.

3.3. VALIDATION AND UNCERTAINTIES

Linear regression analysis was used for validating the methodologies employed, since it allows defining the relationship between two variables, X and Y. For each subject (or experimental unit), knowing both X and Y, the best straight line through the data has to be found. The goal of linear regression is to adjust the values of slope and intercept to find the line that best predicts Y from X. More precisely, the goal of regression is to minimize the sum of the squares of the vertical distances of the points from the line.

3.3.1. Uncertainties estimation strategy

For soil thickness maps, the linear regression between the RW limit depth predicted using the MASW methodology and from the validation penetrometric data set at the location of the penetrometric soundings showed a significant correlation (R^2=0.6255). This means that ~62% of the real RW limit depth should be explained using this linear regression obtained using the MASW methodology. This constitutes a consistency check of the method.

The effective clay content map deduced from hyperspectral data was compared with the interpolated laboratory values map of the total clay mineral concentration, obtained using the Inverse Distance Weighting algorithm, for validation. The general trend shows a good agreement between the predicted and the observed clay distribution tendency. A reliability test of the procedure can be performed using linear regression between hyperspectral-predicted clay values and their correspondent clay mineral content, observed in the sampling data set. Correlation between the two variables is satisfactory, with a determination coefficient of R^2=0.599; as a consequence, our model can explain about 60% of actual clay content in the top level of soils in the study area. The accuracy of clay content estimation and clay maps could be improved through the use of 1414 and 1914 nm clay absorption peaks, which show better correlations with the laboratory dataset, with respect to the used one.

Clay content estimated from EMI measurements may be compared with ground truth measurements for validation. Lower R^2 values of the relationships relating measured clay content to estimated clay content are generally observed when values of soil electrical conductivity are not corrected for soil water content (i.e., when the Rhoades model is not applied and clay content is directly related to measured soil electrical conductivity), which indicates the importance of accounting for this correction. For the 0-10 cm layer, better agreement between clay content estimates and ground truth measurements is observed for the wet soil conditions (R^2=0.27) compared with the dry conditions (R^2=0.18). This would be explained by the more contrasted patterns of soil electrical conductivity over the area under wet than under dry conditions. Nevertheless, in both cases, very poor agreement is observed between measurements and estimated clay content at the location of the anthropogenic soil along the eastern limit of the field. The very different nature of this soil compared with the rest of the studied field may

explain these observations, the values of the model's empirical parameters being soil-specific. Regarding the 20-30 cm layer, poor agreement is found between clay content estimations and measurements, especially within the north-east part of the field where large overestimations of clay content are observed. Such low R^2 values and the sometimes large discrepancies found between ground truth and estimates may at least partly arise from the fact that EMI measurements integrate a large volume of soil while clay content was determined from relatively small samples characterising a rather thin (10 cm thick) layer of soil. Furthermore, the rather low clay content at the study site associated with its narrow range of spatial variation over the investigated area also limit the accuracy of the estimations of this soil property from soil electrical conductivity measurements.

The stone content measured on soil cores (validation dataset) was compared to the stone content measured by the model. Except at some locations in the anthropogenic part, the estimation is rather satisfying, but the model slightly overestimates the real stone content. The Root Mean Square Error has been calculated for the whole calibration dataset, except data from the anthropogenic part: it was equal to 9.7%. An estimation of the stone content at about 10% can be considered as a good estimation, with a precision of the same order of magnitude as visual estimations in the field by a pedologist. The methodology can then be used extensively to determine the stone content at the scale of a parcel. Improvements would consist in better measurements of the stone content on the calibration points, by taking into account larger volumes of soil that would be more representative of the real stone content.

We have 30 ground truth points in different depths to validate the GPR derived soil water content. Only 3 points of them correspond to the dry conditions and the remaining points were sampled during the wet conditions. Therefore, we are only able to validate the wet derived water content estimates. To quantify the error between the 0-10 cm depth ground-truth soil moisture and both GPR derived water content, we used the nearest neighbor of GPR data to the location of ground-truth data. Then linear regressions were used to analyze the relationships between the two variables. The RMSE is about 3.6% for off-ground GPR and 4.9% for on-ground related to the best fit line. Discrepancies between ground-truth measurements and GPR soil water content estimates may arise for several reasons. First, a part of the differences would result from the contrasted characterization depths and scales of GPR compared with ground-truth measurements (100 cm^3 cylinders). Furthermore, while GPR surveys were completed within about one hour and a half, soil core sampling over the entire field took one complete day and soil water content is likely to vary during that time as a result of infiltration, especially within the hours following a rain event as it is the case for this field campaign. Moreover, infiltration is likely to evolve spatially as a result of spatial variation of topography and clay content over the field area.

3.3.2. Spatialization

About spatial error in soil thickness maps, we established the map of the difference between RW limit depth from MASW and from the validation penetrometric data set. It allows the analyzing of the spatial variations of the misfit between the prediction and the validation data set. The prediction error varies between ~-20 cm and ~20 cm. The

isolated positive error spots can be considered as singular values, probably due to a bad picking of the RW limit horizon on the seismic section. For those points, it would be judicious to reiterate one part of the processing workflow. On the other side, we observe that the negative error cloud is concentrated in the area of lowest RW limit depth (eastern area). On a matter of fact, very low soil depth (lower than ~50 to 70 cm) constitute a limitation in terms of resolution for the seismic method.

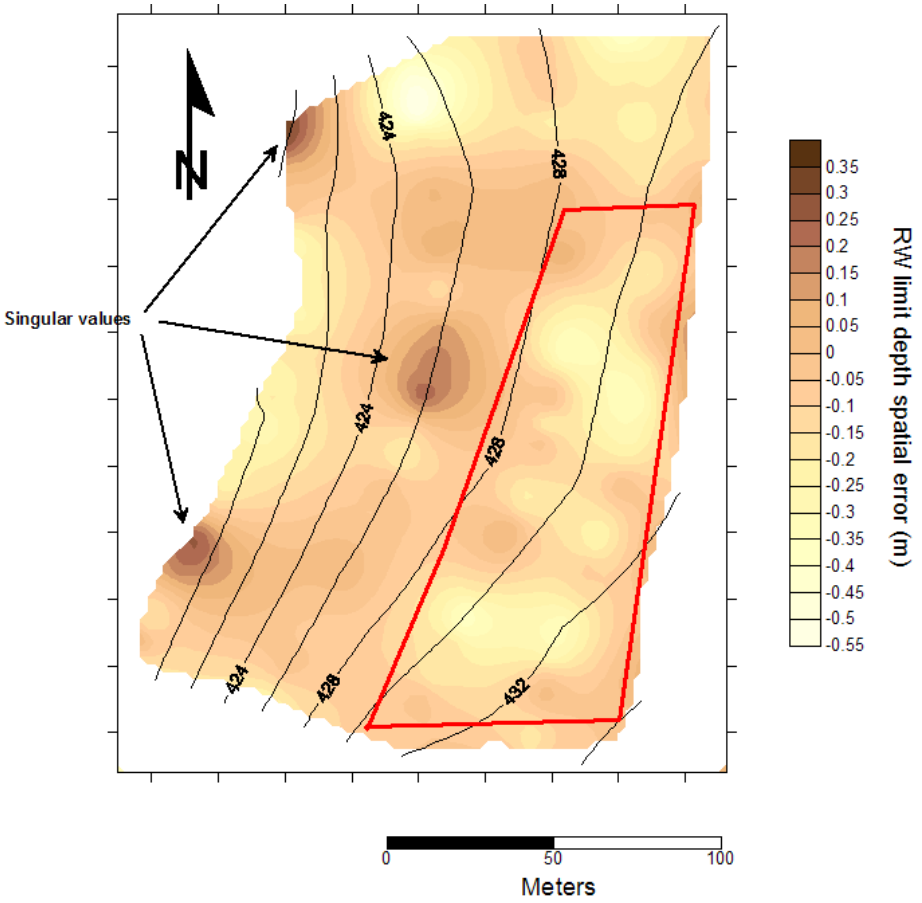


Figure 14 : Map of the RW limit depth spatial error obtained calculating the difference between RW limit depth from MASW and from the penetrometric validation data set

4. Results and discussion

4.1. PROPERTIES MAPS

The geophysical data were interpreted in terms of different thematic maps such as: soil thickness, stone content, water content, effective clay content and C content.

First order maps were obtained by a simple inversion of geophysical signals into soil properties whereas second order maps (such as clay content and C stock) were computed using several first order maps that are combined to produce more accurate information or a soil property of a high level.

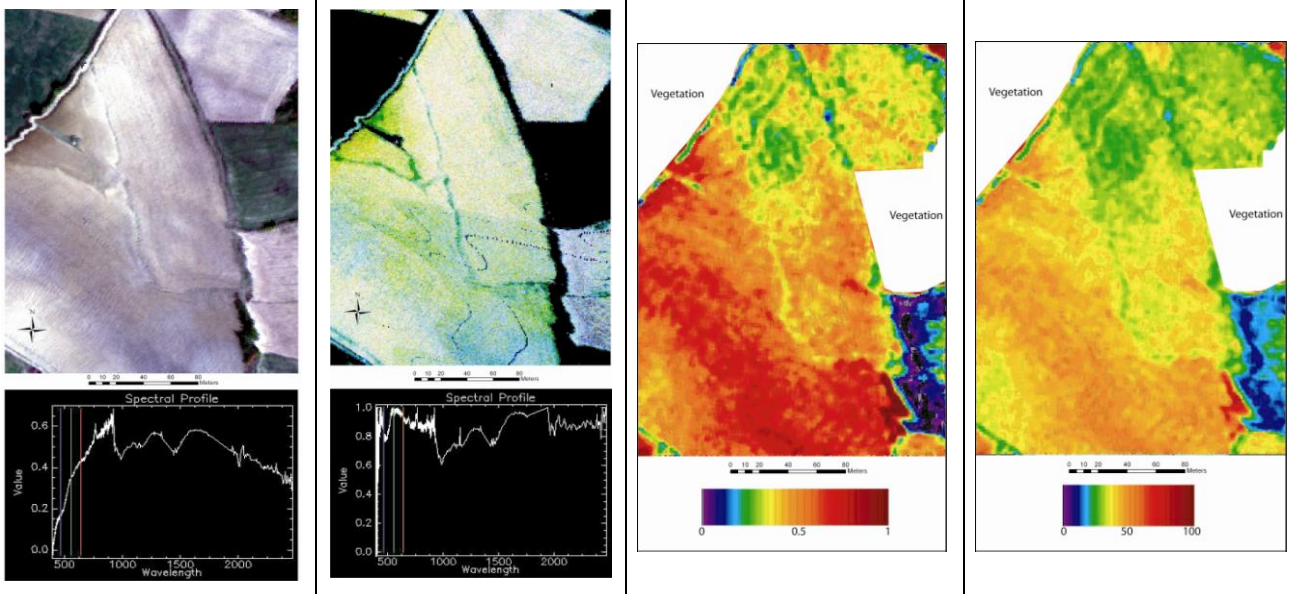
4.1.1. Clay content map

- Effective clay content map

The procedure described hereafter allows obtaining a map of clay content of bare topsoils from airborne hyperspectral images, using the continuum removal technique (Clark and Roush, 1984). This procedure is based on the fact that the depth of an absorption feature is strongly related to the abundance of the absorbing material. Continuum removal normalizes reflectance spectra with the aim of allowing a direct comparison among absorption features from a common baseline, minimizing the effect of different scales or observation conditions and assuming that no other no other material has strong absorption features around that specific wavelength.

After continuum removal the absorption peak depth at a certain wavelength is calculated and related to the atomic group responsible for the spectral feature. Clay content and mineralogy influence the short wave infrared portion of the spectrum (1300-2500 nm), but only the peak at 2210 nm can be detected in airborne sensors spectra, which are affected by atmospheric absorption bands.

Data acquired with Hyper SIM-GA sensor form Selex Galileo were processed in order to obtain geo-referenced, calibrated and atmospherically corrected SWIR cubes. Clay absorption peak depth at 2210 nm from ASD indoor spectra, was correlated with total clay content obtained from laboratory analysis, thus obtaining a calibration line. After continuum removal, the absorption peak depth was calculated at 2210 nm, for every pixel of the image, using a dedicated IDL routine. Then, the laboratory relation was used to invert the hyperspectral images into clay content maps. As a matter of fact, as demonstrated by Lagacherie et al. (2008), laboratory-calibrated functions can be applied at map scale to hyperspectral images to map soil properties. This processing workflow represents a summary of the methodology developed in the report D1.3 and based on the different experiments included in D1.1-2 and 2.1-2.3 and is further explained in the Figure 15 hereafter. Elaborations on images were performed using ENVI software (ITT VIS, Boulder, CO).



Atmospheric corrected SIM-GA image	Step 1) continuum removed image	Step 2) map of absorption peak depths	Step 3) map of percentage clay content

Figure 15 : From acquisition to inversion hyperspectral data

The map of the step 3 shows in false colors the spatial distribution of clay mineral content, resulting from manipulation on hyperspectral images, with orange-red corresponding to higher percentages and green-blue corresponding to lower percentages. A mask was applied to the neighbouring grassland. Lower values are concentrated in the northern part of the parcel, which is topographically more elevated while an increasing trend towards the south (i.e.: parallel to the flow direction, towards the bottom) can be observed.

- Clay content: second order soil map

Soil electrical conductivity of low-salinity soils is mainly affected by soil water content and clay content. Rhoades et al. (1976) proposed the following empirical model to relate soil electrical conductivity to soil physico-chemical properties:

$$\sigma = (a\theta^2 + b\theta)\sigma_w + \sigma_s ,$$

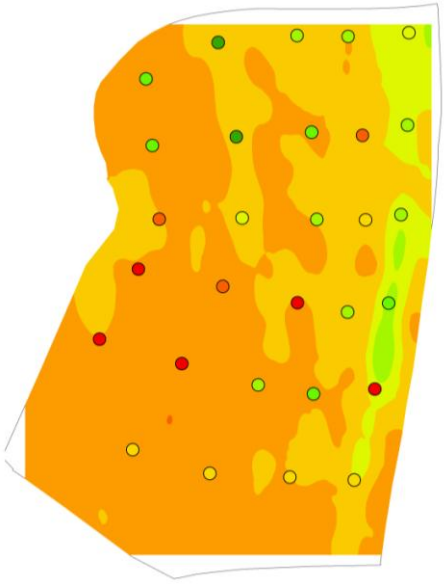
where σ is the bulk soil electrical conductivity ($S.m^{-1}$), θ is the soil water content ($m^3.m^{-3}$), σ_w is the soil solution electrical conductivity ($S.m^{-1}$), σ_s is the electrical conductivity of dry soil ($S.m^{-1}$), and *a* and *b* are soil specific empirical parameters.

In this equation, σ_s may be expressed as a function of the soil clay content. We used this model to estimate clay content from EMI measurements considering the empirical parameter values (*a*=1.382, *b*=-0.093) found by Rhoades et al. (1976) for a soil comparable to that of the main soil unit observed at the study site and assuming σ_w=

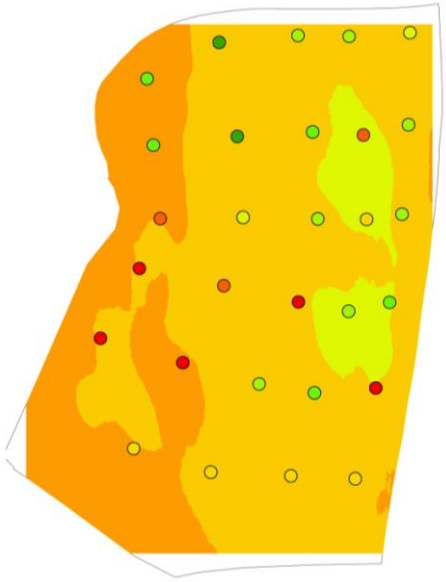
0.05 S.m^{-1}. Water content estimates from GPR measurements were used to remove the effect of θ on measured soil electrical conductivity and provide estimations of σ_s. For a subsample of the ground truth points (0-10 cm and 20-30 cm clay content measured from soil samples collected along the EMI and GPR transects), relationships were then established between σ_s estimates and clay content measurements. Finally, these calibration relationships were applied to estimate soil clay content at each EMI measurement point and these estimations were compared with the complete ground truth measurement data set for validation.

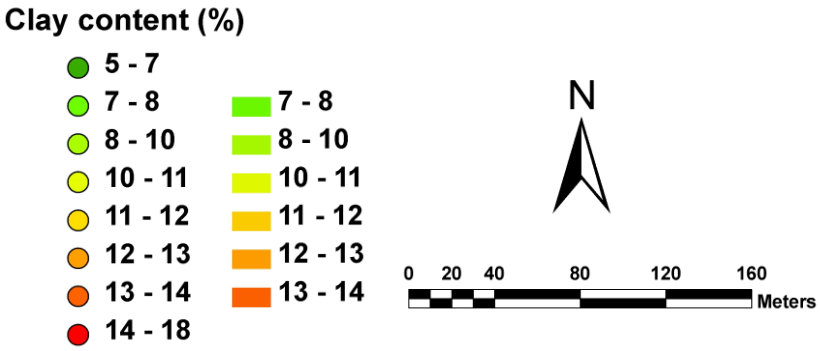
For clay content within the 0-10 cm soil layer, these analyses were performed using horizontal dipoles measurements as they present high sensitivity to the soil surface properties. Data sets from both the first (dry conditions) and second (wet conditions) measurement days were considered in order to compare clay content estimates using data from contrasted soil water content conditions. For clay content below the plough layer (20-30 cm), vertical dipole measurements were used as this configuration presents higher sensitivity to the deeper soil layers, only one set of EMI data (second day of measurements) is available in this case. For each data set, clay content was determined with and without correcting electrical conductivity values for water content (using the Rhoades model), in order to investigate the effect of this correction on clay content estimates. The resulting maps are presented below, in Figure 16.

With correction for water content

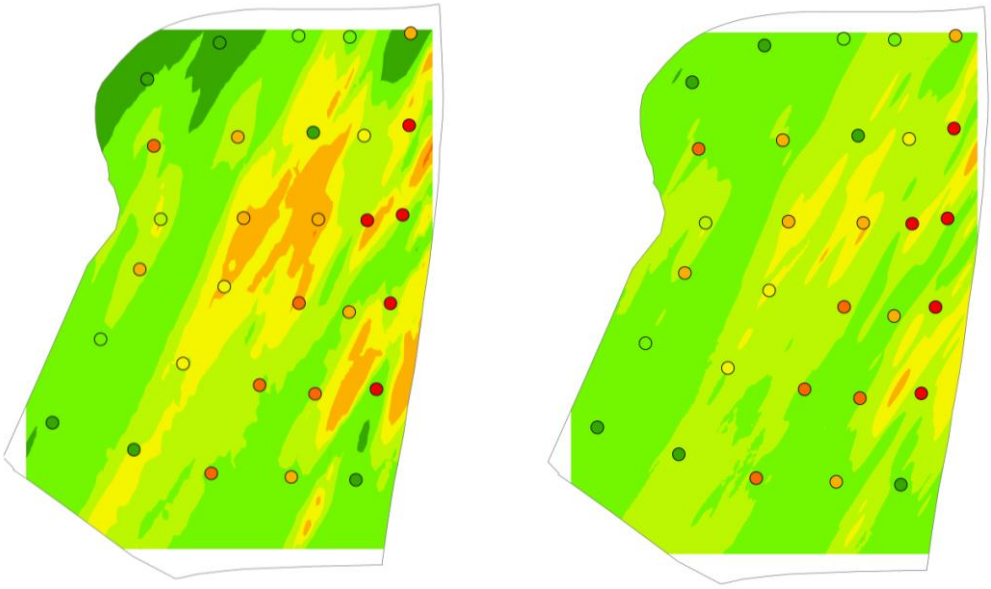


Without correction for water content

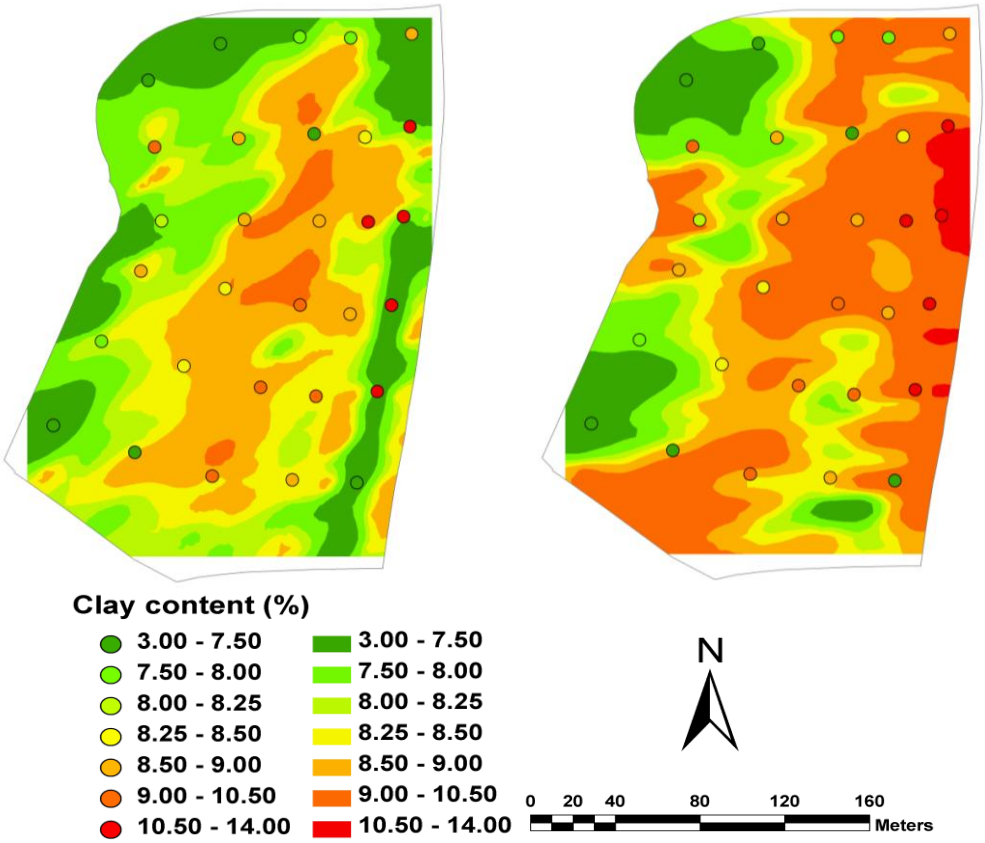




Clay content 0-10 cm, dry conditions



Clay content 0-10 cm, wet conditions



Clay content 20-30 cm, wet conditions

Figure 16 : Kriged maps of clay content estimates from EMI and GPR measurements

4.1.2. Water content map

GPR surveys were carried out using both far-field (off-ground) and near-field (on-ground) systems. Depending on the configuration, different processing procedures were used to retrieve soil water content from the radar data.

We used the Lambot et al. (2004, 2006) full-wave inversion method for far-field radar measurements. In that case, the GPR antenna effects are filtered out by complex linear transfer functions determined by antenna calibration in the laboratory and the filtered frequency-domain signal is converted to the time-domain using the Fourier transform. Then, focusing on the surface wave reflection, the full-wave radar model is inverted to retrieve the soil surface dielectric permittivity. Finally, the soil dielectric permittivity is converted to volumetric water content using Topp's (1980) equation.

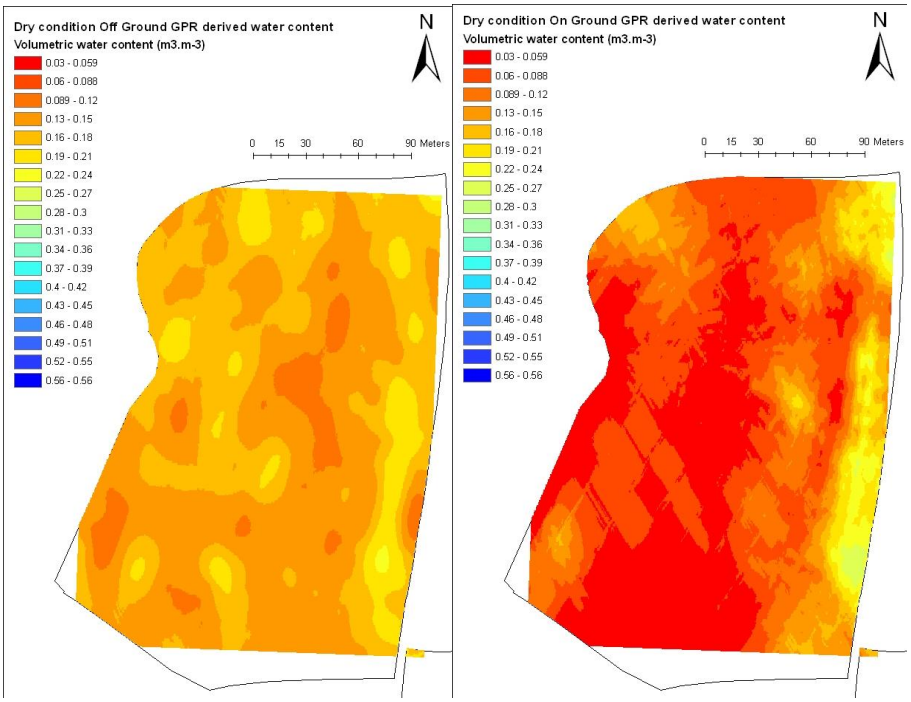
For on-ground GPR, the direct ground wave (DGW) method is used. The single trace analysis (STA) for DGW allows retrieving the soil dielectric permittivity in the shallow soil layer (less than 0.5 m), depending on the antenna center frequency and soil

moisture (**Erreur ! Source du renvoi introuvable.**). In order to retrieve the soil dielectric permittivity, the propagation time from the transmitting antenna (Tx) to the receiving one (Rx) should be calculated. In this case, the Tx-Rx offset is fixed and the ground propagation time is directly related to soil dielectric permittivity and, consequently, to soil water content:

$$\varepsilon_r = \left(c.\frac{\Delta t}{\Delta x} \right)^2 ,$$

where ε_r is soil relative dielectric permittivity (-), c is the speed of light in vacuum ($m.s^{-1}$), Δt is propagation time (s), and Δx is the Tx-Rx offset (m).

Off- and on-ground GPR surveys were carried out at the Luxembourg site during different weather conditions: dry and wet. Applying the STA method to on-ground measurements and surface-reflection inversion to off-ground measurements allowed to retrieved the soil dielectric permittivity from each trace. Then, the soil volumetric water content estimates from Topp's equation were interpolated over the entire field area using kriging. The Figure 17 shows the soil water content maps from both off- and on-ground GPR data in dry and wet conditions. It is worth noting that the off-ground derived map represents only the surface moisture (top ~2-3 cm). The characterization depth for the on-ground map is difficult to specify accurately as it depends on both the operating center frequency and soil moisture. In our case, the characterization depth is expected to be around 30 cm for the dry soil and around 10 cm for the wet conditions.



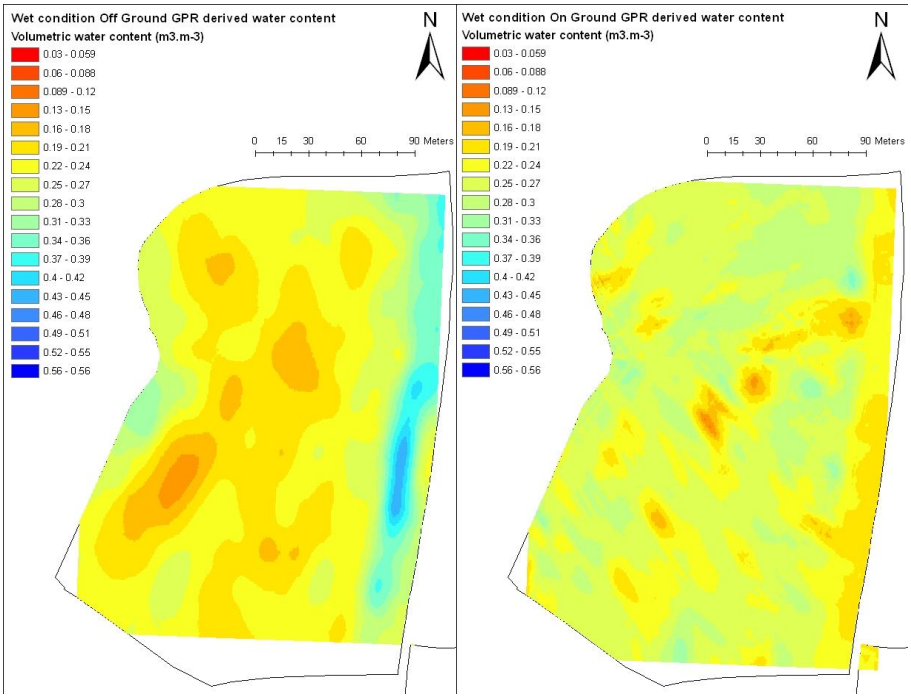


Figure 17 : Soil water content maps derived from off-ground (left) and on-ground (right) GPR data for dry (top) and wet (bottom) conditions at the Luxembourg site.

4.1.3. Stone content map

Input data were electrical resistivity measurements from the three arrays of the ARP. From the first array data, a standard deviation map has been calculated by the following method: each pixel is of 50 x 50 cm in size, and its value is the standard deviation of the electrical resistivity values calculated in a circle centred on the pixel with a radius of 5 m (see paragraph 2.2.1, page 29 in D3.2).

The stone content has been measured on soil cores sampled by a driller over the first 30 cm (see the locations of the sampling in paragraph 2.1, page 17 in D3.2.). The dataset was splitted in two parts: four samples were selected randomly (except in the North-East anthropogenic part) and used as a calibration set.

A linear relationship was determined between the standard deviation of the electrical resistivity and the stone content. As expected, a higher stone content leads to the higher variability in electrical resistivity. To check the stability of the relationship, several random datasets of 4 samples were selected and analysed with the standard deviation of the electrical resistivity. The slope of the relationship was quite stable, equal to about 0.65. The quality of the relationship could be improved by taking into account more numerous samples. We have here decided to use the maximum data for the validation, and, as a consequence, the minimum data for the calibration.

4.1.4. C content map

For this map, we stated that the soil organic carbon content decreases with depth. The spectral models produced the best results for the PSR technique applied to the spectra covering the VIR and SWIR (400-2500 nm). The spectra were calibrated and validated against the C content in the upper 20 cm of the soil. This corresponds to a homogenous C content in the ploughlayer. We excluded one outlier in the validation, as the spectra of the sample point close to the field border was influenced by the neighboring grassland. The PSR on the VNIR-SWIR proved to be the best combination of technique and spectral range.

Stevens et al. (2010) carried out an independent validation of the regional scale spectral models. The models are somewhat less accurate than the field scale model with an RMSE ranging from 4.5 to 5.4 g C kg^{-1} and an RPD from 1.31 to 1.75 (See Figure 18).

In general separate models for each soil type or agricultural region perform best for the full spectral range.

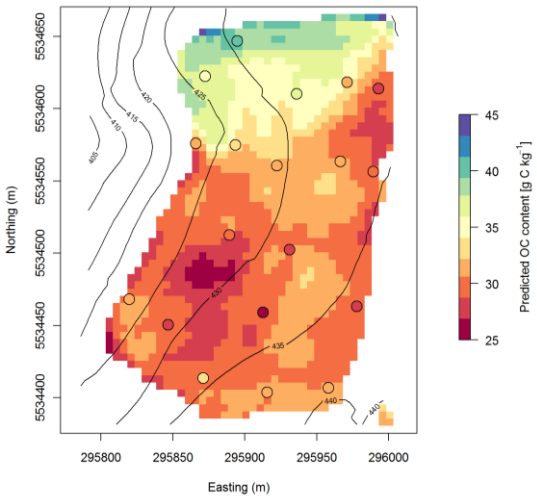


Figure 18 : Map of the C content in the plough layer (0-20 cm).

4.1.5. Soil thickness map

The MASW seismic method allowed obtaining a velocity model at a given point because of the dispersive character of surface waves. But in some case, there is a contraindication for the use of the MASW methodology. Hard soils, non-tabular media particularly don't allow to compute well resolved dispersion diagrams. In such cases, a second processing workflow based on the inversion of P-waves first-time arrivals inversion was used (the example of Luxembourg test site). This method allowed there to retrieve 2D P-waves velocity (Vp) models.

Seismic experiments on Luxembourg site led to realize 5 profiles covering around 5 ha for almost 200 seismic shots. In addition, 30 penetrometer soudings were performed on the same area. The soundings were divided into 2 groups with the purpose of (i) calibrating the Vp isovalue for the soil/bedrock limit horizonation (15 soundings) and (ii) validating the obtain soil depth map using this methodology (15 soundings). In this case, the soil/bedrock limit was defined as the boundary between subhorizontal schistosity red material (~50 to 90 cm) and horizontal schistosity white material (from ~1m) (RW limit). On a matter of fact, the change is the schistosity constitutes a great mechanical contrast which strongly influence surface-waves. The RW limit is identified as a Qd step around 90 cm on the whole penetrometric data set.

When interpolating the soil depth data points obtained from the MASW method, a map of the soil depth was derived. This map was afterwards compared to the validation dataset, i.e., RW limit depth obtained from the 2nd group of penetrometric data, to estimate the a posteriori uncertainty related to the methodology (Figure 19).

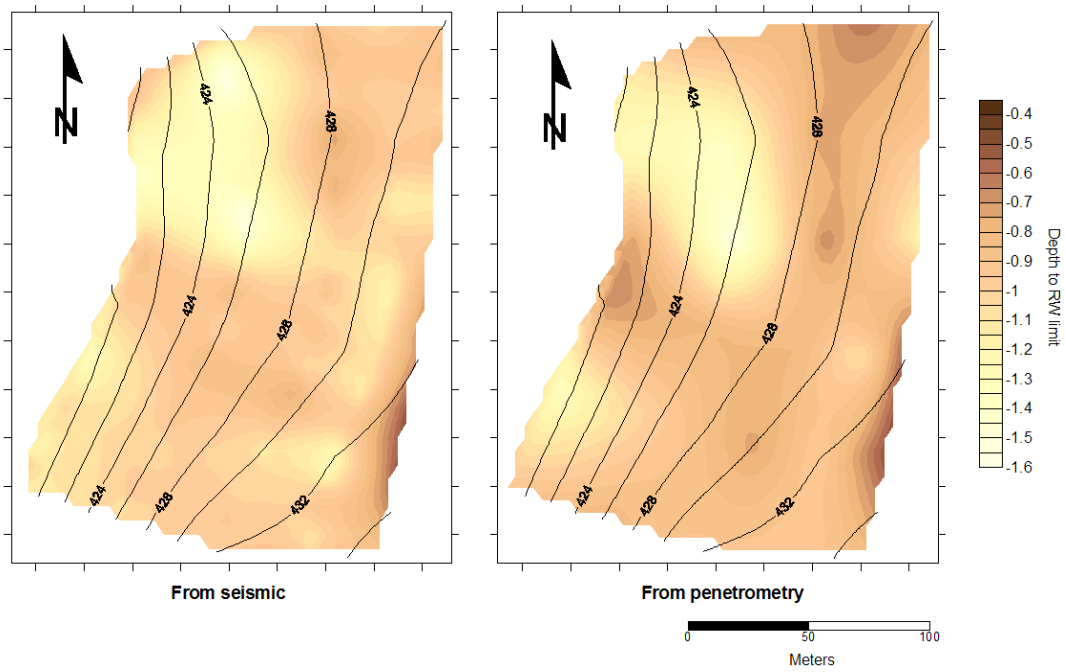


Figure 19 : Comparison between the soil depth obtained from MASW (left) and from the validation penetrometric data set (right).

4.2. TECHNICAL MATURITY AND ECONOMPIC ASPECTS

Testing of the product concepts has been approached by employing a form of marketing research technique, called choice modelling. This typically involves a sample of people, who are expected to make use of a specific good, being asked a series of questions about their preferences for alternative versions of this good. Respondents' choices of their preferred alternatives demonstrated their willingness to trade-off one

attribute against another. Since the alternatives was monetary (i.e. price), it was possible to estimate respondents' willingness to pay (WTP).

4.2.1. Economic evaluation

The survey was administered via email to about a thousand individuals, 166 of whom chose to take part.

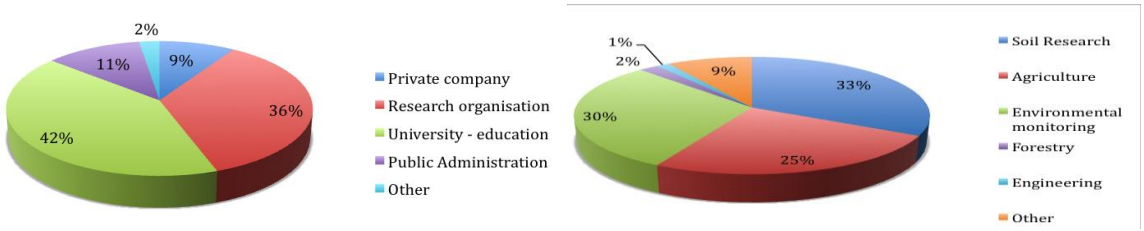


Figure 20: Type of organisation and intended use

The survey began by asking questions on respondents' line of work and on their intended use of the mapping tool. Figure 20 shows the percentage of respondents falling in the various categories chosen for each question.

These figures provide a picture of who are the main target end-users and what is their intended use of the mapping tool. As these figures show, the overwhelming majority of the potential end users come from the research arena, be it a university (42%) or another type of research institute (36%). The third largest group of potential end-users represents public administration entities (11%), followed by the private sector (9%), such as agribusiness companies and consultancies. In terms of the intended use of the DSM, the bulk of the responses are divided roughly equally amongst pure soil research (33%), agriculture (25%) and environmental monitoring (30%).

Attribute	WTP	Attribute	WTP
Map Resolution	183€	Water Content	198€
Soil Depth	157€	Clay Content	264€
Bulk Density	210€	Soil Degradation Indicators	811€
Carbon content	263€		

Table 1: WTP for the several features of the DIGISOIL mapping tool

Table 1 presents the results of the statistical analysis of the responses to the choice experiment part of the survey. It shows the values of the estimated parameters of the respondents' hypothesized utility function.

The way to interpret the above numbers is this: each number represents the average WTP for obtaining an extra level of the respective attribute. For instance, for the

measurement of carbon content, the average respondent would be willing to pay 263€ for obtaining a low-accuracy measurement. In order to have a measurement of high accuracy for the carbon content respondents would be willing to pay 789€ (3 times 263€). The same applies for the other features of the DIGISOIL mapping tool, such as map resolution and inclusion of degradation indicators.

Another point consists to assess the economic potential of the DIGISOIL mapping tool by comparing the costs of producing the various maps with the economic benefits that these maps can confer to the end users. In order to do so, however, the simple enumeration of the aforementioned costs will not suffice. Instead, they need to be interpreted and put in the right context by making certain points and assumptions that will make them comparable to the estimated benefits.

- The estimated WTP is on a per-hectare basis.

- The area covered by the ground-based techniques in Luxemburg is about 6 hectares. The area covered by the airborne hypespectral operation in Mugello is about 130 hectares but the mapped area and the area where soil sampling and analysis was carried out is just over 5 hectares.

- Capital (physical equipment & software) depreciates, and according to the depreciation rate applied, different capital life spans are assumed. An optimistic depreciation rate is 10% for physical capital, implying a time span of 10 years. Computers and software, however, depreciate considerably faster, as newer and more advanced versions spring up regularly, rendering existing ones increasingly obsolete. Thus, with an optimistic 20% rate, software and computer life span is estimated to be 5 years. A maximum of 20 operations can be performed each year.

On the basis of the above hypotheses, a summary of costs per-hectare for each method is reported in Table 2. On the basis of these values a WTP compliance analysis is tried in order to perform a “commercial maturity” evaluation of DIGISOIL technologies and final products (soil parameter maps).

- In the upper part of the table, the WTP/ha values, as estimated and described in D4.1, are reported both for low quality and high quality maps in terms of map resolution and accuracy.

- In the middle part of the table, the different technology costs and instrument configurations for different maps are enhanced. For each kind of map the total cost for the required instrument configuration is also reported.

- The lower part of the table shows the results of the compliance (maturity) analysis for low quality and high quality maps in terms of WTP/Cost (%).

	WTP/ha						
	Map resolution	Soil Depth	Bulk Density	Carbon Content	Water Content	Clay Content	Degradation Indices
Low quality	183	157	210	263	198	264	-
High quality	549	471	630	789	594	792	(811)

Technology	*Cost/ha*							
SISMIC	740		X	X				
GEOEL.	1590			X		X	X	
GPR/EMI	1035		X		X	X		
HYPER	270				X		X	
Total	*3635*							
		Technology Cost/ha/Map	1775	2330	1305	2625	1860	(9895 = total cost/ha)
		Low Quality Map Maturity	8.8%	9.0%	20.1%	7.5%	14.1%	
		High Quality Map Maturity	26.5%	27.0%	60.4%	22.6%	42.5%	

Table 2. "Commercial matury" analysis based on cost results and estimated WTP

The “commercial gap” which results from WTP/cost compliance analysis means that further business development on digital soil mapping, as well as further technology developments and economic studies are needed to bridge this gap up to the full commercial maturity of DIGISOIL products. This should be particularly the object of future project more oriented toward operational prototypes.

5. Conclusions

In order to assess and prevent soil degradation, and to benefit from the different ecological, economic, and historical functions of the soil in a sustainable way, there is an obvious need for high-resolution, accurate maps of soil properties. The core objective of the project was to explore and exploit new capabilities of advanced geophysical technologies for answering this societal demand. Several geophysical techniques were carried out in the project are based on positive experiences in the domain, and promise to fulfill the objectives of the project. Electrical and GPR measurements, hyperspectral imagery, and more innovative methods like seismic methods were tested and technically adapted to soil properties mapping. An important output of the project concerned the economic and maturity estimation of such technologies, based on the real experiments carried out on the two DIGISOIL test sites.

Technical maturity and economic analyses were detailed in D4.1 and D4.2, and were performed on the basis of willingness to pay (WTP) and real cost analysis methods. The main result of this work is the commercial maturity estimation for low and high quality maps where:

- The unit cost/ha ranking from (low to high cost) of studied geophysical methods is: 1) Hyperspectral (270€), 2) Seismic (740€), 3) GPR/EMI (1035€), 4) Geoelectric (1590€)

- The maturity ranking of soil maps is 1) Carbon Content (60%), 2) Clay Content (42%), than Bulk Density, Water Content and Soil Depth in the range (23-27%)

- The commercial gap range from 77% to 40% for high quality Water Content map and Carbon Content maps respectively.

The “commercial gap” which results from WTP/cost compliance analysis means that further business development on digital soil mapping, as well as further technology developments and economic studies are needed to bridge this gap up to the full commercial maturity of DIGISOIL products. This should be particularly the object of future project more oriented toward operational prototypes.

6. References

Archie, G.E., 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. Petroleum Transactions of AIME, 146, 54–62.

Atzberger, C. 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models, Remote Sensing of Environment 93 (1–2): 53–67

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

Banton, O., M.K. Seguin, and M.A. Cimon. 1997. Mapping field-scale physical properties of soil with electrical resistivity. Soil Sci. Soc. Am. J., 61, 1010-1017.

Ben-Dor E., Y. Inbar and Y. Chen, 1997 The reflectance spectra of organic matter in the visible near infrared and short wave infrared region (400-2,500nm) during a control decomposition process. Remote Sensing of Environment 61,1-15.

Ben-Dor E., K. Patkin, A. Banin and A . Karnieli 2002 Mapping of several soil properties using DAIS-7915 hyperspectral scanner data. A case study over clayey soils in Israel. International Journal of Remote Sensing 23:1043-1062.

Besson A., Cousin I., Richard G., and Boizard H., 2004. Structural heterogeneity characterization of the soil tilled layers by a 2D electrical resistivity prospecting. Soil and Tillage Research, 79, 239-249.

Binley, A., Cassiani, G., Middleton, R. and Winship, P., 2002. Vadose zone flow model parameterisation using cross-borehole radar and resistivity imaging. Journal of Hydrology, 267,3-4, 147-159.

Cassiani, G. and Binley, A., 2005. Modeling unsaturated flow in a layered formation under quasi-steady state conditions using geophysical data constraints. Advances in Water Resources, 28, 5, 467-477.

Chang, C., D.A. Laird, M.J. Mausbach and C.R. Hurburgh. 2001. Near-infrared reflectance spectroscopy-Principal Components Regression analyses of soil properties. Soil Science Society America Journal, 65: 480-490.

Cierniewski, J. and A. Karnieli. 2002. Virtual surfaces simulating the bidirectional reflectance of semi-arid soils. International Journal of Remote Sensing, 23: 4019–4037.

Cohen, M.J., J.P. Prenger and W.F. DeBusk. 2005. Visible-Near Infrared reflectance spectroscopy for rapid, nondestructive assessment of wetland soil quality. Journal of Environemental Quality, 34: 1422-1434.

COM232, 2006. Proposal for a Directive of the European Parliament and of the Council establishing a framework for the protection of soil and amending Directive. COM/2006/0232, http://eur-lex.europa.eu. 30p.

Cousin, I., Besson, A., Bourennane, H., Pasquier, C., Nicoullaud, B., King, D. and Richard, G., 2009. From spatial-continuous electrical resistivity measurements to the soil hydraulic functioning at the field scale. C. R. Acad. Sci., 341, 10-11, 859-867.

Dalal, R.C. and R.J. Henry. 1986. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. Soil Science Society America Journal, 50: 120-123.

Daniel, K.W., N.K. Tripathi and K. Honda. 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.

Daniels, D.J., 2004. Ground Penetrating Radar, 2nd Edition. The Inst. Electrical Eng., London.

Darvishzadeh, R., A. Skidmore, M. Schlerf and C. Atzberger. 2008. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland, Remote Sensing of Environment 112 (5): 2592–2604

Eckelmann, W., Baritz, R., Bialousz, S., Bielek, P., Carre, F., Houskova, B., Jones, R.J.A., Kibblewhite, M.G., Kozak, J., Le Bas, C., Toth, G., Varallyay, G., Yli Halla, M., Zupan, M.,

2006. Common criteria for risk area identification according to soil threats. European Soil Bureau Research Report n°20, EUR 22185 EN, 94 pp. Office for Official Publications of the European Communities, Luxembourg.

Fidêncio, P.H., R.J. Poppi and J.C. de Andrade. 2002a. Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy. Analytica Chimica Acta, 453: 125-134

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

Geladi, P and BR Kowalski. 1986. Partial Least-Squares Regressions: a tutorial. Analytica Chimica Acta, 185: 1-17.

Giao, P.H., Chung, S.G., Kim, D.Y., and Tanaka, H., 2003. Electric imaging and laboratory resistivity testing for geotechnical investigation of Pusan clay deposit. Journal of Applied Geophysics, 52, 4, 157–175.

Grandjean, G., Malet, J.P., Bitri, A., and Meric O., 2007. Geophysical data fusion by fuzzy logic for imaging mechanical behaviour of mudslides. Bull. Soc. Geol. France, 177, 2, 133-143.

Hermann, R.B., 1987. Computer programs in seismology. Saint-Luis University, USA.

Islam, K., B. Singh and A. McBratney. 2003. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared spectroscopy. Australian Journal of Soil Research, 41: 1101-1114.

Kalinski, R.J. and W.E. Kelly. 1993. Estimating Water Content of Soils From Electrical Resistivity. ASTM Geotechnical Testing Journal. 16, 3, 323-329.

Grandjean, G., 2006a. A seismic multi-approach method for characterizing contaminated sites. J. Applied Geophys., 58, 87-98.

Grandjean G., 2006b. Imaging subsurface objects by seismic P-wave tomography: numerical and experimental validations. Near Surface Geophysics, 275-283.

Grandjean G. and Bitri A., 2006. 2M-SASW: inversion of local Rayleigh wave dispersion in laterally heterogeneous subsurfaces: application to Super-Sauze landslide (France). Near Surface Geophysics,, 367-375.

Grandjean, G., Paillou, P., Dubois-Fernandez, P., August-Bernex, T., Baghdadi, N. and Achache, J., 2001. Subsurface structures detection by combining L-band polarimetric SAR and GPR data: Example of the Pyla Dune (France). Ieee Transactions on Geoscience and Remote Sensing, 39, 6, 1245-1258.

Jacquemoud, S., C. Bacour, H. Poilve and J.-P. Frangi. 2000. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode, Remote Sensing of Environment 74 (3): 471–481

Lambot, S., Slob, E.C., van den Bosch, I., Stockbroeckx, B., Scheers, B. and Vanclooster, M., 2004a. Estimating soil electric properties from monostatic ground-penetrating radar signal inversion in the frequency domain. Water Resources Research, 40, W04205, doi:10.1029/2003WR002095.

Lambot, S., Slob, E.C., van den Bosch, I., Stockbroeckx, B. and Vanclooster, M., 2004b. Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties. IEEE Transactions on Geoscience and Remote Sensing, 42, 2555-2568.

Lambot, S., Weihermüller, L., Huisman, J.A., Vereecken, H., Vanclooster, M. and Slob, E.C., 2006. Analysis of air-launched ground-penetrating radar techniques to measure the soil surface water content. Water Resources Research, 42: W11403, doi:10.1029/2006WR005097.

Lanz, E., Mauer, H., Green, A.G., 1998. Refraction tomography over a buried waste disposal site. Geophysics 63, 1414-.1433.

Loke, M.H. and Barker, R.D., 1996. Rapid least-squares inversion of apparent resistivity pseudosections using a quasi-Newton method. Geophysical Prospecting, 44, 131-152.

Michot D., Benderitter Y., Dorigny A., Nicoullaud B., King D., Tabbagh A., 2003. Spatial and temporal monitoring of soil water content with an irrigated corn crop cover using surface electrical resistivity tomography. Water Resources Research, 39, 5, 11381158.

Malley, D., Martin, P., and Ben Dor, E., 2004. Application in analysis of soils. Pages 729-784In: Craig. R.; Windham R.; Workman J. (Editors), Near Infrared Spectroscopy in Agriculture. A Three Society Monograph (ASA, SSSA, CSSA), 729-784.

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

Mc Mechan, G.A., Yeldin, M.J., 1981, Analysis of dispersive waves by wave-field transformation, Geophysics, 46, pp.869-874.

Nakashima, Y., Zhou, H. and Sato, M., 2001. Estimation of groundwater level by GPR in an area with multiple ambiguous reflections. Journal of Applied Geophysics, 47, 241-249.

Panissod, C., Dabas, M., Jolivet, A., Tabbagh., A., 1997. A novel mobile multipole system (MUCEP) for shallow (0–3m) geoelectrical investigation: the 'Vol-de-canards' array. Geophysical Prospecting 45, 983–1002.

Park, C.B., Miller, R.D., Xia, J., Ivanov, J., 2000. Multichannel seismic surface-wave methods for geotechnical applications. Proc. of the First Int. Conf. on the App. of Geophys. Methodologies to Transportation Facilities and Infrastructure, St. Louis, December 11–15.

Tabbagh, A., Dabas, M., Hesse, A., Panissod, C., 2000. Soil resistivity: a non-invasive tool to map soil structure horizonation. Geoderma 97, 393–404.

Toth, G., 2008. Soil quality in European Union. In Toth et al. Eds. Threats to Soil Quality in Europe. JRC Report EUR 23438 EN, 162p.

Reeves III, J.B. and S.R. Delwiche. 2003. SAS partial least squares regression for analysis of spectroscopic data. Journal of Near Infrared Spectroscopy, 11: 415-431.

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

Stevens, A., van Wesemael, B., Vandenschrieck, G., Tychon, B., Touré, S., 2006. Detection of carbon stock change in agricultural soils using spectroscopic techniques. Soil Science Society of America Journal 70, 844–850.

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

Sturtevant, K. A., Baker, G. S., Snyder, C., Kopczynski, S., 2004. Hydrogeophysical characterization of bedrock fracture orientations using azimuthal seismic refraction tomography. AGU, H23A-1122.

Tarantola, A., 1987, Inverse problem theory. Elsevier Science Publishing Co., Inc.

Whiting, M.L., L. Li and S.L. Ustin. 2004. Predicting water content using Gaussian model on soil spectra. Remote Sensing of Environment, 89: 535-552.

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

Scientific and Technical Centre
RNSC Division
3, avenue Claude-Guillemin - BP 36009
45060 Orléans Cedex 2 – France – Tel.: +33 (0)2 38 64 34 34