

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".



Measuring and processing protocols description

FP7 – DIGISOIL Project Deliverable 1.3

FP7-DIGISOIL-D1.3

October 2009

S. Lambot, F. Andre (UCL/FZJ)

With the collaboration of

G. Grandjean, K. Samyn (BRGM), B. Van Wasemael, A. Stevens (UCL)

L. Chiarantini (GAV), I. Cousin (INRA)

Checked by:

Name: S. Lambot

Date: 15/03/2009



Approved by:

Name: G. Grandjean

Date: 20/03/2009



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".



Keywords: geophysical parameters, soil characteristics, data fusion.

In bibliography, this report should be cited as follows:

S. Lambot, F. Andre, G. Grandjean, K. Samyn, B. Van Wasemael, A. Stevens, L. Chiarantini, I. Cousin, 2009. Measuring and processing protocols description. Report N°BRGM/FP7-DIGISOIL-D1.3, 60 pages.

© BRGM, 2009. No part of this document may be reproduced without the prior permission of BRGM.

Synopsis

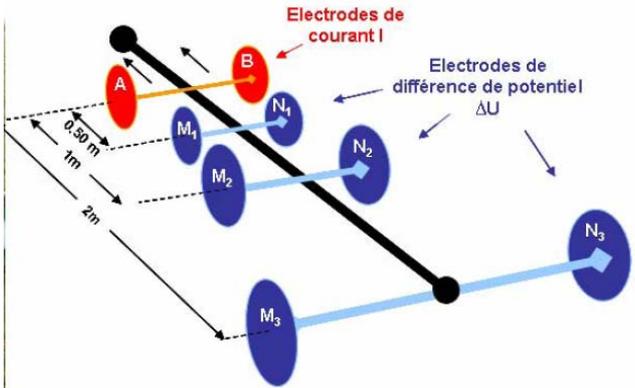
The present deliverable concerns the last task of the DIGISOIL's WP1. Possible and adapted processing or inversion techniques for geophysical data are studied.

A state of the art on sensor technologies, general theory related to inverse problem is first given. Then, for the 5 methods developed in DIGISOIL (GPR, EMI, Geoelectric hyperspectral and seismic), interpretation, limitations and field examples are compiled in synthetic tables format.

Contents

1. Geoelectric method	7
2. Seismic method	11
3. GPR method	21
4. EM method	29
5. Hyperspectral method	39
6. Conclusions	49

1. Geoelectric method

GEOELECTRIC	
Sensor description	
<p>The MUCEP device was initially developed by CNRS Garchy (Panissod et al., 1997; Dabas et al., 2001). The system is composed by a SYSCAL acquisition station 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, measurements 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 by 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.</p>	
	<p>Photo of the MUCEP system The data acquisition station is located inside the vehicle; The first two wheels carry the injection electrodes while the last pairs carry measurements channels for 3 different spacings</p>
	<p>Sketch of the MUCEP system Identification of the different electrodes (current & potential) and spacings between them used in the field acquisitions</p>

	<p>The SYSCAL acquisition system</p> <ul style="list-style-type: none"> - Automatic injection ranging, stacking (microprocessor controlled). - Current 2.5A Maximum - Voltage 1,000V standard(2,000V peak to peak), up to 1,500V with external converters. - Power 250 Watt, or up to 1,200W with external converter. - Pulse Duration 0.15, 0.25, 0.50, 1.0, 2.0, 4.0, and 8.0 seconds. - Current Precision 0.2% typical - Automatic SP correction - Simultaneous measurement of voltage and current.
<p><i>Table 1 : Description of the MUCEP system. The electrodes part and the SYSCAL acquisition system.</i></p>	

GEOELECTRIC	
Signal processing	
Data corrections	
<p>To compare electrical resistivity measurements at different date, data have to be corrected from temperature effect. The equation of Keller and Frishknecht (1996) can be used efficiently:</p>	
$\rho_{T_{ref}} = \rho_m \left[1 + c(T_M - T_{ref}) \right] \quad (1)$	
<p>Where $\rho_{T_{ref}}$ is the corrected electrical resistivity at the Tref reference temperature, ρ_{T_m} is the observed electrical resistivity at the Tm temperature and c is a coefficient mostly equal to 2%.</p>	
Inversion theory	
<p>Since we operate the MuCEP device with three interelectrodes spacings, we interpret 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 was developed at the University of Paris VI (France) and that uses the Levenberg-Marquadt optimisation algorithm (Cousin et al., 2009). The inverse process is constrained by: (i) a conceptual model of the electrical resistivity distribution on the vertical plan; and (ii) an estimation of layer's thickness. Another approaches can also be used to invert for resistivity profiles, like IX1D (Interpex) or RES2DINV (Loke and Barker, 1996). Basics of electrical imagery need to deal with back projection: let us consider the scalar potential U; the continuity equation for the current density can be expressed as following:</p>	
$\text{div}(\sigma \text{grad}U) = 0 \quad (2)$	
<p>where σ is the electrical conductivity. Equation (2) together with the modified complete electrode model equations are discretized numerically in the usual way, so that approximate values of electrode voltages for the approximate element conductivity vector σ can be calculated, assuming the constant approximation of the conductivity on each of all elements. The forward calculation yields an estimation of the electric potential field in the interior of the volume under certain Neumann and Dirichlet boundary conditions. Finite elements of finite differences in 1, 2 or 3D can be exploited for this forward problem with current sources. Image reconstruction of is an inverse problem, which is usually presented as minimizing the suitable objective function $\psi(\sigma)$ relative to σ. To minimize this objective function we can use a deterministic approach based on the Least Squares method. Due to the ill-posed nature of the problem, regularization has to be</p>	

used as shown in equation (3):

$$\psi(\sigma) = \frac{1}{2} \sum \|U_M - U_C\|^2 + \alpha \|L\sigma\|^2 \quad (2)$$

Where U_M and U_C are measured and calculated potentials; α and L are regularization parameter and matrix respectively.

References

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.

Dabas, M., A. Tabbagh, and D. Boisgontier. 2001. Multi-depth continuous electrical profiling for characterization of in-field variability. In E. C. O. P. Agriculture. Montpellier.

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.

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

GEOELECTRIC

Retrieval of the soil properties

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 1:

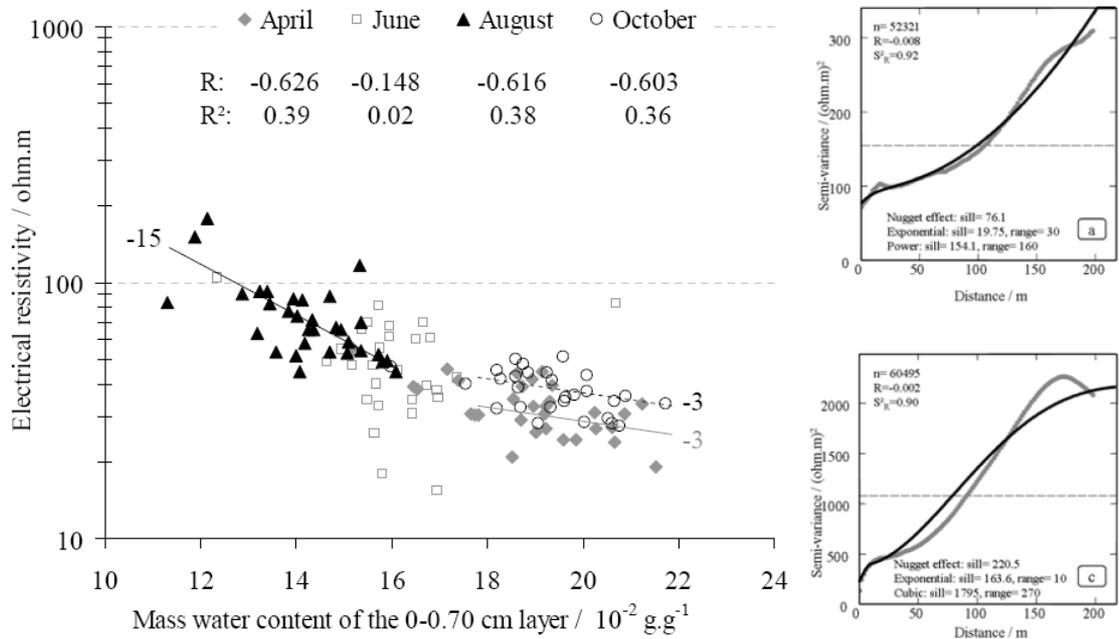


Figure 1 : LEFT: Relationship between the electrical data from the MUCEP Channel 1 and the water content measured in the 0-0.70 m soil layer. RIGHT: Experimental (grey) and theoretical

(black) variograms of the electrical resistivity measured at MUCEP channel 1 with the cross-validation indicator at two seasons. From Besson (2007).

The variography analysis is presented on Figure 1-RIGHT. For each date, the nested models show different basic structures composed of a nugget effect and a spherical structure. The indicators of cross-validation are close to zero for the R mean error, and close to unity for the S^2R ratio of the mean squared error to the krigging variance. That means that whatever the date, the calculated maps of electrical resistivity shows a similar spatial organization.

References

Besson, A., 2007. Analyse de la variabilité spatio-temporelle de la teneur en eau des sols à l'échelle parcellaire par la méthode de résistivité électrique. PhD Thesis, Univ. Orléans, France.

GEOELECTRIC

Operational conditions and limits

The main issue for operating electrical resistivity measurements lies first in the surface soil conditions that can restrict (in cases of high resistivity) the intensity of emitted current in the soil and thus increase the signal to noise ratio. Another point lies in the lack of resolution that could occur when electrode spacing are too sparse or the limitation in the penetration depth when electrode array is too small.

Field testing

In Figure 2, an example of water content mapping is provided. In this study (Besson, 2007) the above-described method is used for going from the electrical resistivity to the water content after a variogram analysis, modelling the spatial correlation between these two quantity, has been carried out.

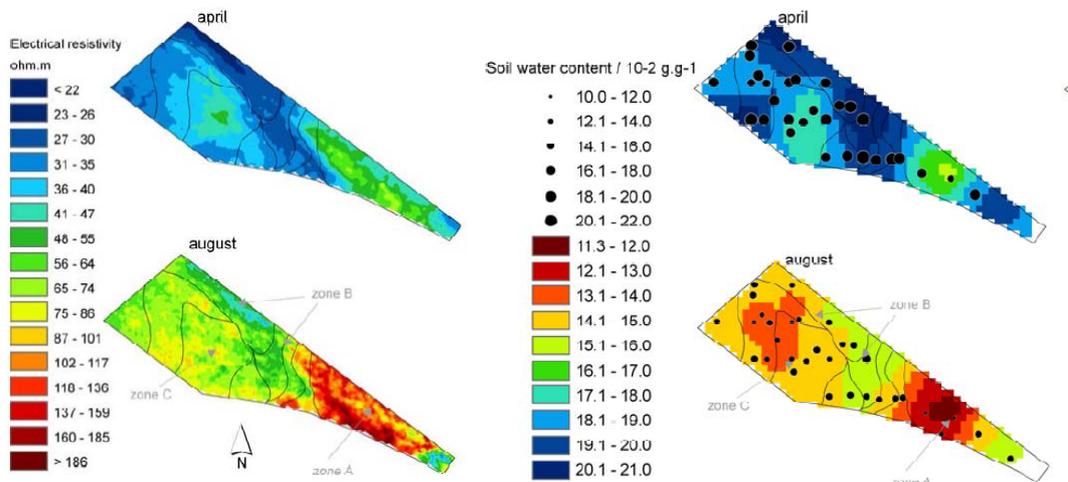


Figure 2 : Maps of the electrical resistivity (left) and water content (right) for a parcel at two different seasons. From Besson (2007).

2. Seismic method

SEISMIC		
Sensor description		
<p>Three seismic sensors are investigated in DIGISOIL: (1) vertical component geophones of 10 Hz central frequency from Geospace, (2) vertical component geophones of 100 Hz central frequency from Mark Products, and (3) tractable gambles of 10 Hz central frequency from Mark Products. The main characteristics of these sensors including those of the acquisition unit are depicted in Table 1. In the same way, three distances between receivers have been tested for this project: 10, 20 and 50 cm. Shots points were systematically set at 50 cm from the first receiver. Optimal spacing that allows a precise velocity picking on the dispersion diagram is 50 cm because the maximum energy can be clearly observed on the whole bandwidth and propagation modes can be distinguished with accuracy. The 10 Hz geophones or gambles offer a large coverage of the whole frequency band and stay in good agreement with the required objectives. The fundamental and first modes are well identified in the dispersion diagram. As this tool has been made available only recently, there are not yet scientific studies which demonstrated its performances. To establish the final seismic device's configuration, let's resume the acquisition system is composed by:</p> <ul style="list-style-type: none"> - A hammer striking on an anvil for generating the source signal; - A seismic antenna of 24 receivers with central frequency of 10 Hz; - A sensor spacing of 50 cm. - A Geometrics GEODE unit for the numerical management of the diverse signals - A PC-based central unit for piloting the acquisition using the Geometrics recording software; 		
	<p>GeodeTM (Geometrics) 24 Channels Bandwidth : 1.75 Hz to 20 KHz Data format : SEG-2 Trigger accuracy: 1/32 of sample interval</p>	
	<p>Geophones (Geospace) 375 Ohms Main frequency: 10 Hz Plugged on tractable device</p>	

	<p>Gamble (Mark Products) Main frequency: 10 Hz</p>
<p><i>Table 1 : Description of the sensors and acquisition unit that are investigated in DIGISOIL. From top to bottom: Geometrics seismograph, 10 Hz geophones and 10 Hz gamble</i></p>	

SEISMIC	
Signal processing	
Dispersion analysis	
<p>The spectral analysis of surface waves (SASW) method is an <i>in situ</i> 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: (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. 2D wavefield transform method is used to determine experimental dispersion curve. Wavefield transformation methods involve a 2D wavefield transform in which the data are transformed from space–time domain into a more convenient domain. The dispersion curve, i.e., the plot of phase velocity versus wavelength (or frequency), is associated with the peaks in the transformed spectrum. The frequency–slowness (f-p) transform involves first a Radon or t-p transform on the data, followed by a one-dimensional (1D) Fourier transform along the t direction. McMechan and Yedlin (1981) described a technique to obtain phase velocity dispersion from an array of seismic traces, which one is used for this study. They proposed first performing 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</p>	
$A(\omega, r_n) e^{i\phi(\omega)_n}$	
<p>One possible p-ω stack of N traces at different distances from the same source is defined by the relation</p>	
$F(p, \omega) = \sum_{n=1}^N C(\omega, r_1, r_n)^{-1} A(\omega, r_n) e^{i\phi_n} e^{i\omega r_n}$	
<p>Where</p>	
$C(\omega) = A(\omega, r_1, r_n) e^{i\phi_1} \sqrt{\frac{r_n}{r_1}}$	
<p>Searching for the maxima of</p>	
$ F(p, \omega) $	
<p>yields the possible dispersion curves. Since there are N distances, the maximum value of the</p>	

quantity $|F(p, \varpi)|$ should be equal to N. In the realistic case of multimode surface waves, the last equation will not yield a maximum independent of the amplitude spectrum of the other modes. The stack value will typically be largest for one mode and smaller for others. In the case of a uniform half-space, the dispersion curves extracted from amplitude spectrum in the frequency-velocity domain should exhibit a nearly constant wave velocity with increasing frequency. Figure 1a presents a typical shape of dispersion curve performed with a layered medium. The corresponding medium profile is presented in Figure 1b. The portions of the curve associated with the sensitivity to the various layers of the medium in Figure 1b are illustrated in an approximate way across the top of the dispersion curve.

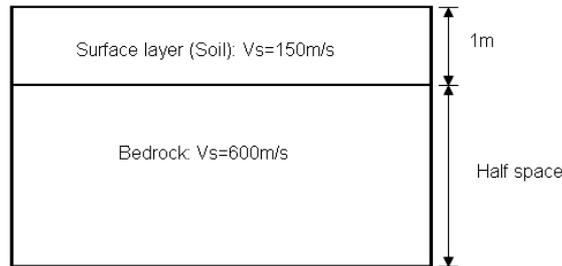
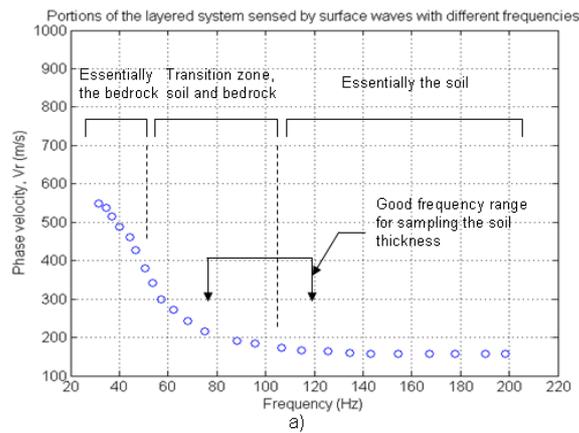


Figure 1 : Typical dispersion curve showing approximate ranges in frequencies associated with different layers: (a) dispersion curve; (b) idealized medium profile.

Inversion theory

The inverse problem aims at finding the parameters which characterize in the best way the medium (here the Vs and thicknesses of layers). The real dispersion curves, compared to synthetic curves, create residues of phase velocity, which constitute the y data of the inverse problem. 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 non linear relation (Tarantola, on 1987). This non linear problem can however be linearized to an iterative problem defined by:

$$\left[(Ax - y)^t (Ax - y) + \sigma^2 x^t x \right] = MIN$$

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:

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

The matrix of the *a posteriori* covariance gives an estimation of the trust interval of the model:

$$C = V(\Lambda^2 + \sigma^2 I)^{-1} \Lambda^2 (\Lambda^2 + \sigma^2 I)^{-1} V^t$$

In the same way, the resolution matrix is obtained with:

$$R = V(\Lambda^2 + \sigma^2 I)^{-1} \Lambda^2 V^t$$

A weighting factor *W* was introduced in order to control the variation of the parameters of the model for each iteration. The inversion process is implemented according to Herrmann (1987).

References

- Herrmann, R.B., 1987. Computer programs in seismology. Saint-Luis University, USA.
 Mc Mechan, G.A., Yeldin, M.J., 1981, Analysis of dispersive waves by wave-field transformation, Geophysics, 46, pp.869-874.
 Tarantola, A., 1987, Inverse problem theory. Elsevier Science Publishing Co., Inc.

SEISMIC

Retrieval of the soil properties

Near surface methods using elastic wave propagation is conducted at frequencies that vary between a few Hz to a few kHz. In this frequency range, the wavelength in soil ranges between tens of centimetres to tens of meters, therefore the wavelength is much greater than the grain size and the seismic wave propagate without perturbation through the soil mass. There are three important propagation modes in the near surface: longitudinal propagation (P-wave), transverse propagation (S-wave) and retrograde elliptical Rayleigh wave (R-wave). The shear modulus of the soil G_{soil} , depends on the skeleton shear stiffness, $G_{soil} = G_{sk}$, and is not affected by the bulk stiffness of the pore fluid. For this reason the shear waves are preferred for the characterisation of the near surface deposits. The S-wave velocity V_s is:

$$V_s = \sqrt{\frac{G_{soil}}{\rho_{soil}}}$$

where ρ is the mass density of the soil. The shear modulus is determined by the state of stress, the degree of cementation and by processes that alter inter-particle contacts (capillarity forces, electrical forces). Shear wave velocity can be lower than 20 m/s Figure 2 for soil near suspension to skeleton transition, and can reach 300 m/s to 500 m/s at the depth 30 m. For the unsaturated soils, the bulk stiffness of the fluid is very low, the bulk and shear moduli of the soil mass are those of the soil skeleton. Poisson's ratio is lower than 0.15. For any degree of saturation, the velocity of shear waves is determined by cementation, state of effective stress, capillary forces in silty or clayey soils etc. The propagation of longitudinal P-wave is proportional to constrain modulus M and the mass density ρ

$$V_p = \sqrt{\frac{M_{soil}}{\rho_{soil}}} = \sqrt{\frac{B_{soil} + \frac{4}{3} G_{soil}}{\rho_{soil}}}$$

where B_{soil} is the bulk modulus and G_{soil} is shear modulus of the soil. For the saturated soil, the P wave velocity varies between 1500 m/s and about 2000 m/s, depending on porosity. The Poisson's ratio approaches to 0.5. For the unsaturated soil the velocity of the P wave is about 1.4 to 1.6 times higher than shears velocity. The free soil surface promotes the formation of Rayleigh R-wave. The velocity of R-wave V_R is related to the S wave and the P wave velocities and can be estimated as (Santamaria et al. 2005)

$$V_r \approx \frac{0.874 + 1.117\sigma}{1 + \sigma} V_s$$

where σ is Poisson's ratio. For unsaturated soils $V_r \approx 0.9V_s$. The R-wave may be used to measure V_s avoiding the need for borehole or probs. The dispersive nature of R-wave propagation in vertically heterogeneous medium forms the basis of surface wave testing (Stokoe et al. 1994, Soco et al. 2004, Matthews et al. 1996,). For determining bulk density profiles from V_s ones, the use of elastic parameters is mandatory. With the SASW technique, it is useful to obtain a 2D imagery of the density, which can highlight the strong or the weak zone density in an area. Several ways may be performed to obtain the density from the V_s measurements. It is useful to test all of these ways, so as to determine the most accurate method.

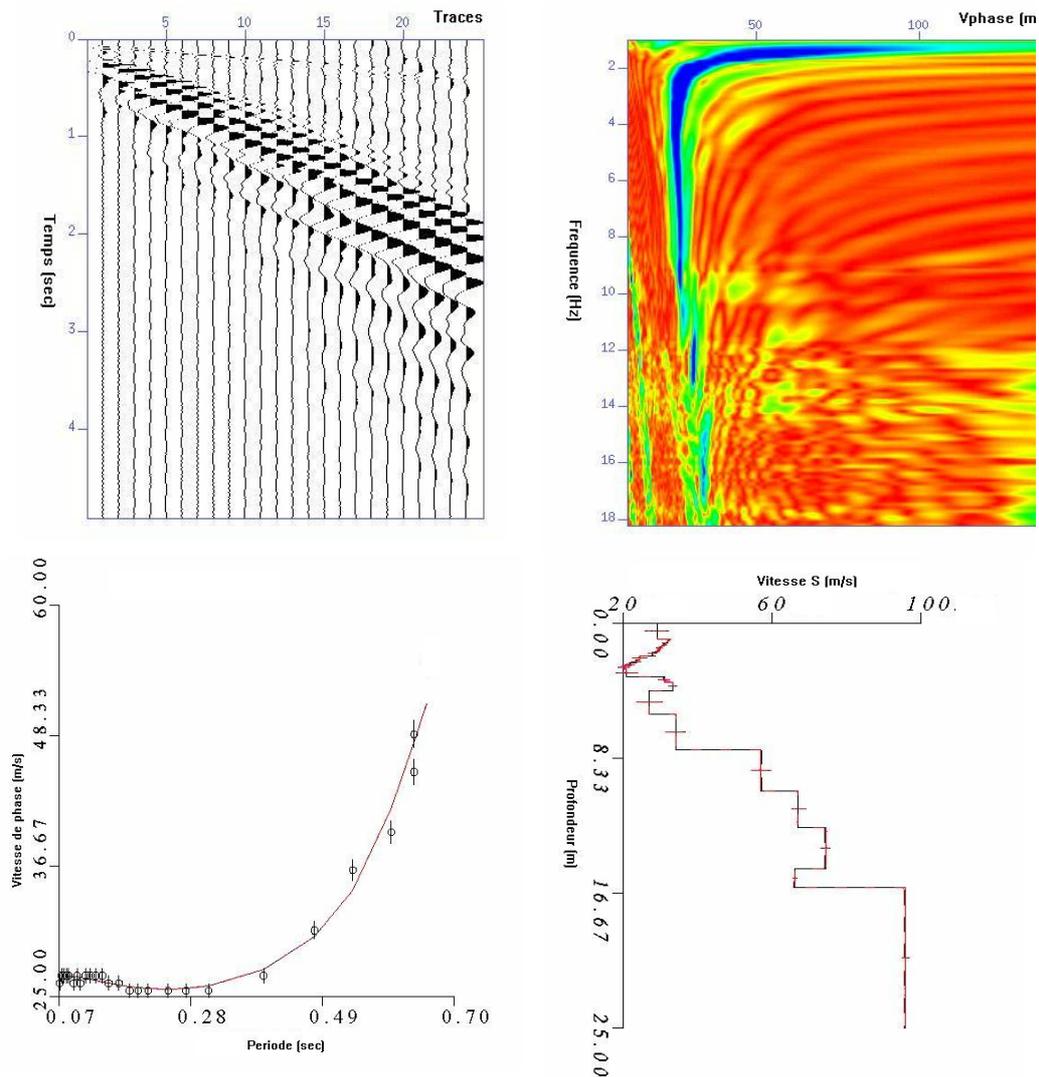


Figure 2 :a-Example of short record with obvious low velocity Rayleigh wave (b)- Phase velocity plot obtained by interception time ray parameter transform c- Experimental and theoretical phase velocity fitting by linear inversion d-Depth variation of shear velocity.

Some empirical relationships can be found between bulk density ρ and shear wave velocity V_s .

For saturated soil materials, Burns and Mayne (1996) expressed this relation:

$$\rho = 0.701 (V_s)^{0.227} (\sigma'_{v0})^{-0.057}$$

where σ'_0 is the effective vertical overburden stress in kPa, V_s is in m/s. Another statistical relationship is presented by Mayne et al. (1999) for gravels, sands, silts and clays:

$$\rho = 1 + \frac{1}{0.614 + 58.7 (\log z + 1.095) / V_s}$$

where z is the depth in meters, V_s is in m/s (Figure 3). Some other empirical relations exist between compression wave velocity V_p and either bulk density or porosity. These relations depend on the type of soil, and on the consolidation state. They can be linear or quadratic relations. Then, with using the relation between V_p and V_s with the Poisson coefficient, we obtain a relation between ρ and V_s^1 .

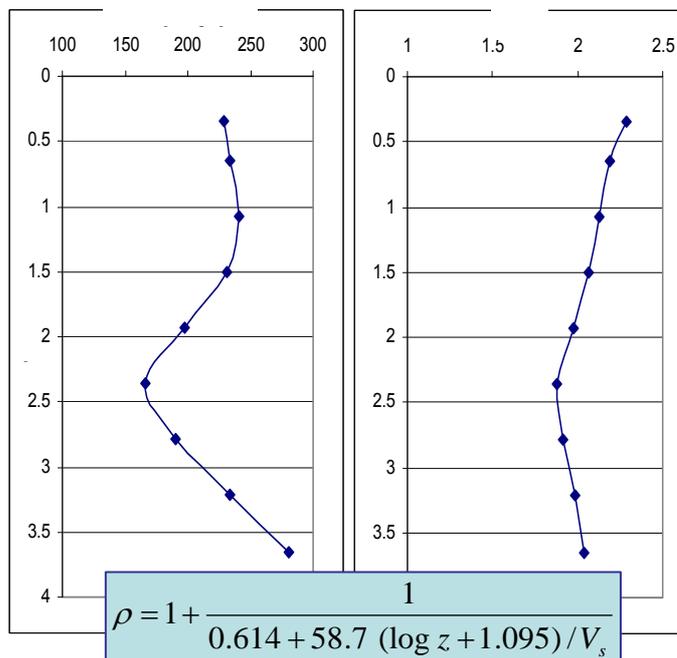


Figure 3: Estimating ρ from V_s values.

Another possibility is to use *situ* geotechnical parameter q_d , which is the dynamic resistance measured with the PANDA penetrometer. Correlation between V_s and q_d is shown on figure 4. Some studies (Gourves and Barjot, 1995) have highlighted correlations between q_d and q_c obtained from the static penetrometer CPT:

$$1 q_d \text{ (MPa)} = 1 \text{ CPT (MPa)}$$

Then some empirical relations are established between maximum shear modulus G_{\max} and q_c , for various types of soils. For the sand (Lunne et al., 1997):

$$G_{\max} = 1634 q_c \left(\frac{q_c}{(\sigma'_v)^{1/2}} \right)^{-0.75}$$

ρ'_v , q_c et G_{\max} in kPa

For the clays (Mayne and Rix, 1993):

$$G_{\max} = 406 (q_c)^{0.695} e^{-1.130}$$

It must be highlighted that these relations correlate a small strain parameter such as G_{\max} with penetration parameter that relates to much larger strains. Finally with the following elastic relation:

$$G_{\max} = \rho V_s^2$$

We can obtain the bulk density ρ^2 . To check the validity of these correlations and to verify if they are adapted to this context, it could be useful to obtain the bulk density by another measurement. Thus, some correlations exist between the dynamic resistance q_d measured by the PANDA penetrometer, and the porosity, if the water content is known, as it is demonstrated by Bernard and Dudoignon (2007). Indeed, Perdok et al. (2002) developed with laboratory tests a relation between the resistance according to the porosity and the water content for clays, and sands, and for a water content comprised between 9 to 20 %. The equation is of the following form:

$$\rho = (1 - n)\rho_s + Sr.n.\rho_f$$

ρ_s and ρ_f correspond respectively to the solid and fluid densities, Sr is the saturation index, n is the porosity. For obtaining the n value, let assumes that:

$$\log(q_d) = a_0 + a_1n + w(a_2 + a_3n)$$

The coefficients a_0 , a_1 , a_2 and a_3 depend on the nature and the structure of the soil. To obtain the parameters adapted to the site, if we are in saturated conditions, the equation becomes:

$$\log(q_d) = a_0 + a_1 \frac{\gamma_s w}{1 + \gamma_s w} + w(a_2 + a_3 \frac{\gamma_s w}{1 + \gamma_s w})$$

So in that case, four values of (w , q_d) are sufficient to establish uniquely this relation.

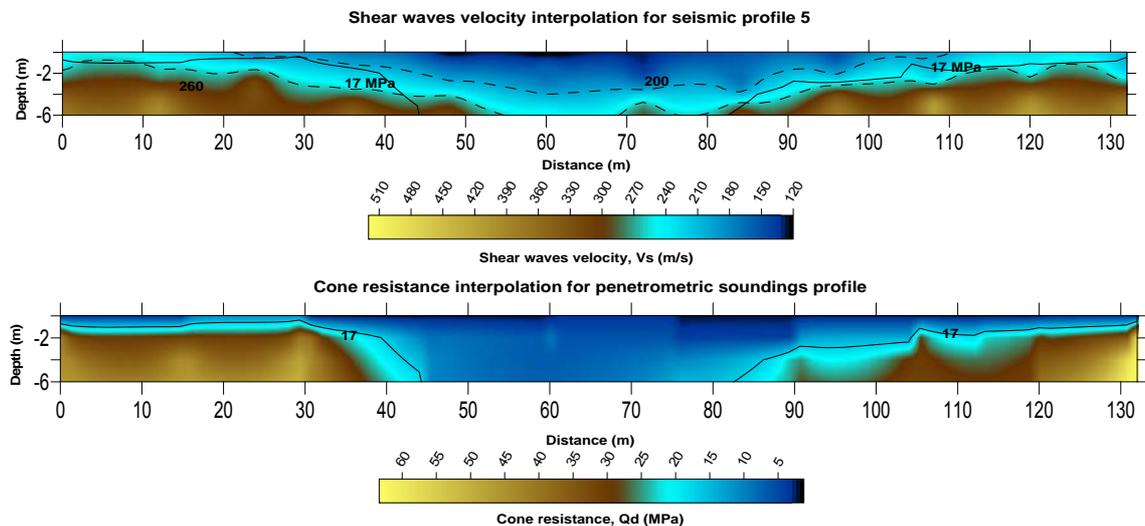


Figure 4: Comparison between shear waves velocity section (top) and cone resistance section from penetrometric measurements (bottom). Horizon of interest on penetrometric soundings can be identified on the seismic profile.

References

Bernard M., Dudoignon P., 2007. Corrélation profils de résistance à la pointe – profils de teneur en eau dans les sols argileux : Marais de Rochefort. 25èmes rencontres de l'AUGC, 23-25 mai 2007, Bordeaux.

Burns, S.E., Mayne, P. W. 1996. Small- and high-strain measurements of in-situ soil properties using the seismic cone penetrometers. TRR 1548 Small-Magnitude Measurements in Geotechnical Engineering. National Academy Press, pp. 81-88.

Gourves R. and Barjot R., 1995. The PANDA ultralight dynamic penetrometer for soil investigation. Proc 11th Euro. Conf. on Soil Mechanics and Foundation Engineering, 28th May – 1st June 1995 Copenhagen.

Lunne P., Robertson P. K., Powell J. J. M., 1997. Cone penetration testing in geotechnical practice. E and FN SPON.

Matthews, M.C., Hope, V.S., Clayton, R.I., 1996. The use of surface waves in the determination of ground stiffness profiles. Proc. Instn. Geotech. Engng., 119, 84-95.

Perdok U. D., Kroesbergen B. and Hoogmoed W. B., 2002. Possibilities for modelling the effect of compression on mechanical and physical properties of various Dutch soil types. Soil and Tillage Research, 65:61-75.

Santamarina J.C, Rinaldi V.A., Fratta D., Klein K.A., Wang Y.H., Cho G.C., Cascante G. A. 2005. Survey of Elastic and electromagnetic properties of near surface soils. In Near Surface Geophysics. Investigation in geophysics No 13.

Socco, L.V and Strobbia, C., 2004. Surface-wave method for near-surface characterization: a tutorial. Near Surface Geophysics, 165-185.

Stokoe, K.H, Wright, G.W., James, A.B., and Jose, M.R., 1994. Characterisation of geotechnical sites by SASW method, in Geophysical characterisation of site, ISSMFE Technical Committee # 10 edited by R.D Woods, Oxford Publishers, New Delhi.

SEISMIC

Operational conditions and limits

Using seismic techniques we expect to produce at the catchment scale the distribution of seismic velocities (P and S-waves) along cross sections reaching 1 or 2 m depth. As the resolution is depending on the frequencies available in the recorded signal we expect to have several cm of resolution in order to identify the soil layering and at least the depth of the bedrock. Following this way of thinking, the main expected progress was to adapt the seismic source and the seismic antenna for working at high frequencies with the capability to be moved quickly; In this way, sensors are coupled and united to a guide for faster installation and this device is towed behind a caterpillars-type vehicle (Figure 5). Data are manually acquired and are simultaneously geo-referenced with a tracking GPS. Besides, extraneous factor like presence of unconsolidated layers (peats, humus etc.) at the surface can affect the measurements by limiting in this case the transmission of high frequencies and in the same way the resolution. An evident advantage of the method lies in the character of properties deduced from the seismic survey which produces the only mechanical waves of the DIGISOIL project. Several other key characteristics of surface waves and surface-wave imaging give strengths to this application. First and probably foremost is the ease with which surface waves can be generated. The relative high-amplitude nature of surface waves (in comparison to body waves) makes their application in areas with elevated levels of mechanical/acoustic noise possible. A half-space is all that is necessary to propagate surface waves. Surface-wave propagation does not require the velocity to increase with depth and/or a contrast at a boundary (i.e., velocity, density, or combination [acoustic impedance]). Conductivity of soils, electrical noise, conductive structures, and buried utilities all represent significant challenges to electrical or EM methods. These have little or no impact on the generation or propagation and generally no influence on the processing or interpretation of surface-wave data. This flexibility in acquisition and insensitivity to environmental noise allow successful use of shear-wave velocity profiling in areas where other geophysical methods may be limited.



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.

Field testing

For the mapping of the soil shear waves velocity, seismic shots are performed along a regular grid for which spacing are determined according to the chosen spatial resolution. Figure 6 shows a zone covered by seismic investigation for one of the DIGISOIL site.

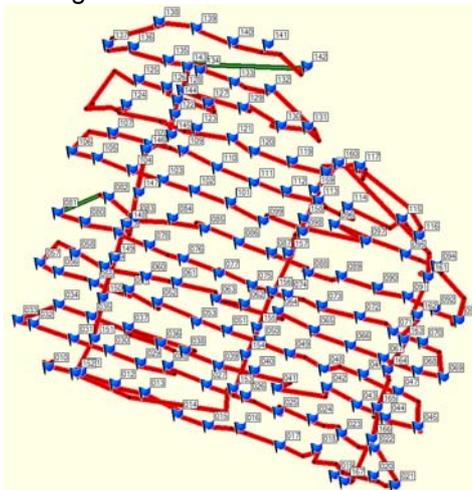


Figure 6: Coverage of a studied zone for seismic investigation.

Through comparison of shear waves velocity of the medium and geotechnical parameters from penetrometric soundings, horizonation of geological limits can be done (Figure 4). Figure 7 shows the interpolation of such boundary in the 3D distribution of seismic velocities collected on the same DIGISOIL test site as Figure 6. Finally, we produce at the catchment scale, a map of the mean shear waves velocity of the soil and a map of the soil thickness.

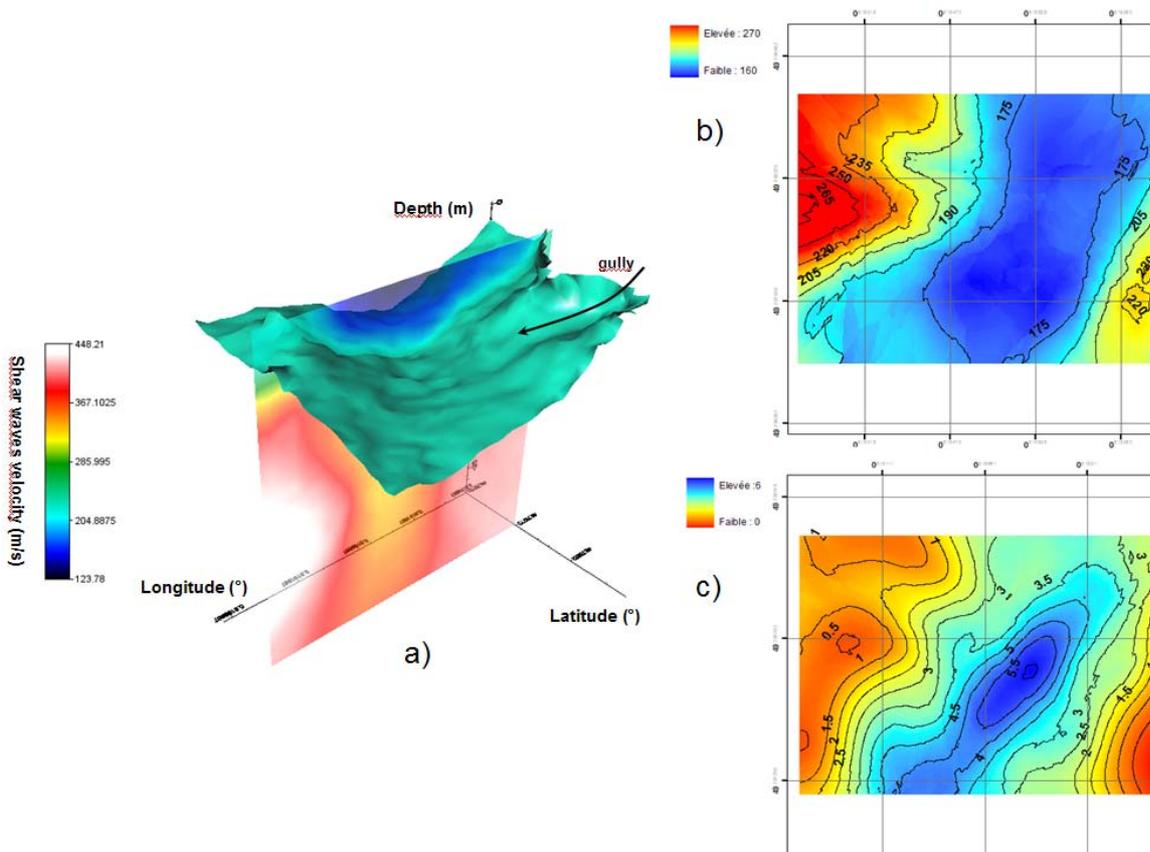


Figure 7 : a) representation of the 230 m/s isosurface (soil-Bedrock limit) in the 3D seismic velocity representation; b) map of the mean shear waves velocity; c) map of the soil thickness obtained from the interpolation of the soil-bedrock horizon.

3. GPR method

GROUND PENETRATING RADAR	
Sensor description	
<p>Two ground penetrating radar (GPR) sensors are investigated in DIGISOIL : (1) a frequency-domain, zero-offset, off-ground GPR prototype set up with vector network analyzer (VNA) technology and an off-ground directive horn antenna, and (2) a time-domain ground-coupled commercial system. The main characteristics of these sensors are depicted in Table 1. 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. The ground-coupled mode of the commercial system permits deeper characterization. Multi-offset measurements can be performed to increase information content in the data. Presently, there are no full-waveform models that are available to fully describe such radars. As a result, simplified models are used, usually based on the straight-ray approximation, for quantitative estimation of the soil properties. Yet, ground-coupled systems are particularly useful for qualitative subsurface imaging with a high resolution.</p>	
	<p>GPR prototype (UCL/FZJ)</p> <ul style="list-style-type: none"> ▪ VNA ZVL from Rohde & Schwarz → SFCW radar ▪ Double ridged horn antenna BBHA 9120F from Schwarzbeck ▪ Antenna dimensions : 96 x 68 x 95 cm ▪ Frequency range : 200-2000 MHz ▪ Mode of operation : zero-offset, at 1.1 m above the ground
	<p>SIR-20 (GSSI)</p> <ul style="list-style-type: none"> ▪ Time domain radar (pulse radar) ▪ Dipole antennas ▪ Antenna dimensions depend on centre frequency ▪ Centre frequency : 400 MHz (typical for soil, but other centre frequencies are available) ▪ Mode of operation: ground-coupled, zero- or multi-offset (2 channels available)
<p><i>Table 1 : Description of the two GPR systems that are investigated in DIGISOIL.</i></p>	

GROUND PENETRATING RADAR	
Signal processing	
Traditional GPR	
<p>We briefly describe GPR techniques which are commonly used with commercial time domain radars to identify surface and depth-dependent soil dielectric permittivity and correlated water content. Soil parameter retrieval is usually performed using ray-tracing based algorithms, thereby implying the plane wave approximation and straight- or curved-rays. As shown below, the soil dielectric permittivity can be retrieved from the determination of the reflection coefficient or from the determination of the wave velocity, which is linked to the dielectric permittivity according to:</p>	
$\varepsilon_r = \left(\frac{c}{v}\right)^2 \quad (1)$	
<p>where c is the wave velocity in free space and v is the wave velocity in the soil. The wave velocity in the soil is determined by measuring the propagation time of the wave (from the radar data) and knowing the travel path length.</p>	
<p>The surface reflection method for soil surface water content estimation</p>	
<p>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:</p>	
$R = \frac{1 - \sqrt{\varepsilon_r}}{1 + \sqrt{\varepsilon_r}} \quad (2)$	
<p>where ε_r is the relative dielectric permittivity of the soil. The soil dielectric permittivity can therefore be derived as:</p>	
$\varepsilon_r = \left(\frac{1 - R}{1 + R}\right)^2 \quad (3)$	
<p>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:</p>	
$\frac{R}{R_{PEC}} = \frac{\frac{E_s}{E_i}}{\frac{E_{s,PEC}}{E_i}} \quad (4)$	
<p>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 (4) reduces to:</p>	

$$R = -\frac{E_s}{E_{s,PEC}} = -\frac{A}{A_{PEC}} \quad (5)$$

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.

The surface ground wave method

This technique applies to ground-coupled GPR systems. A ground-coupled GPR works with source and receiver antennas located at soil-air interface; in this configuration, four fundamental waves can be distinguished and represented by ray paths and wave-fronts (see Figure 1). The waves generated by the transmitting antenna propagate in the upper (air) and lower (soil) media. The ground wave is the signal that travels directly from source to receiving antenna through the soil surface. The ground wave is received at the early times (after the air wave), but its evanescent character makes it hard to detect with large antenna offsets. The soil surface dielectric permittivity can be determined knowing the transmitter-receiver distance and ground wave propagation time. The time delay Δt between the air and ground wave can be estimated through the equation:

$$\Delta t = \frac{d}{c}(\sqrt{\varepsilon_r} - 1) \quad (6)$$

where d is the antenna separation, c is the speed of light (3×10^8 m/s) and ε_{ra} is the relative, apparent soil dielectric permittivity. The depth of influence of GPR ground-wave data for soil moisture estimation has been estimated by various investigators to be between 10–50 cm, depending on acquisition parameters, mainly the operating radar frequency, and soil conditions. When the ground wave has been identified using several measurements with increasing antenna separations, the ground wave velocity can be determined from measurements with a fixed antenna separation (single trace analysis (STA)). However, STA leads to higher uncertainties on the wave speed estimation compared to multiple measurements and the use of multiple receivers for mapping applications is recommended. The limitations of the technique for practical field applications are the required contact between the antennas and the soil, the identification of the ground wave, which may be ambiguous or even impossible in some conditions, and the presence of ambiguous guided waves when near-surface layering is present. In addition, no information on the soil electrical conductivity is provided. All of this may restrict the application of this method for agriculture and other environmental mapping applications.

Vertical dielectric profile determination

A commonly used method to identify the depth dependent dielectric permittivity that governs wave propagation speed is the common midpoint (CMP) method. With this method, stacking velocity fields are extracted from multi-offset radar soundings, i.e., measurements with different antenna separations, at a fixed central location. Wave propagation velocities in the ground can be obtained using the Pythagorean theorem or by tomographic inversion. The dielectric permittivity can then be directly computed and related to the water content. Yet, CMP-derived velocity estimates are generally characterized by low resolution and high uncertainty. In addition, the most important practical disadvantage of CMP is presently that it can not be used for real-time mapping using bistatic radar systems as it requires several measurements, with the antennas in contact with the soil, for a single profile characterization. The use of antenna arrays is a solution for fast surveys. A different approach to reconstruct vertical dielectric profiles is to resort to borehole GPR, where the transmitting and receiving antennas are lowered at different depths in boreholes separated by some known distance. The soil dielectric permittivity is then

determined from the determination of the wave propagation velocity between the boreholes from the measured travel-times. Conducting multi-offset tomographic measurements permits in particular to reconstruct two-dimensional images of the soil dielectric permittivity and correlated water content between the boreholes. The borehole method is not appropriate for field scale characterization ad to cover relatively large areas, but permits high-resolution characterization and monitoring over distances of a few meters.

GPR prototype

The processing procedure that is used is similar for both GPR and EMI. For the GPR prototype, the measured quantity is the frequency-dependent complex ratio, $S_{11}(\omega)$, between the received signal at port 1 of the VNA and the transmitted signal at port 1. This ratio contains both the subsurface response and the antenna effect information. It may be represented by the block diagram shown in Figure 2 and accordingly expressed in the frequency domain by the following equation (Lambot *et al.*, 2004):

$$S_{11}(\omega) = \frac{b(\omega)}{a(\omega)} = H_i(\omega) + \frac{H_t(\omega)G_{xx}^\uparrow(\omega)H_r(\omega)}{1 - H_f(\omega)G_{xx}^\uparrow(\omega)} \quad (7)$$

where $b(\omega)$ and $a(\omega)$ are, respectively, the backscattered and incident waves at the VNA reference calibration plane; $H_i(\omega)$, $H_t(\omega)$, $H_f(\omega)$ and $H_r(\omega)$ are the characteristic antenna transfer functions accounting for the antenna effects and antenna-soil interactions; G_{xx}^\uparrow is the transfer Green's function of the air-subsurface system modelled as a 3-D multilayered medium, each n th layer being considered as homogeneous and characterized by its dielectric permittivity (ϵ_n), electrical conductivity (σ_n) and thickness (h_n). The magnetic permeability is assumed to be constant and equal to the permeability of free space. The characteristic antenna transfer functions can be determined by solving a system of equations as (7) for different model configurations. To this end, we use well defined model configurations, i.e., with the antenna situated at k different heights above a perfect electrical conductor (copper sheet), so that the Green's functions G_{xx}^\uparrow can readily be computed and the corresponding $S_{11}(\omega)$ can be measured in a standard way. The system of equations should be over-determined ($k > 3$) to ensure a well-defined and accurate solution (the equations may not be fully independent for the whole frequency range, depending on the measurement heights). Function $H_f(\omega)$ can also be determined in an independent way, by performing measurements in free space conditions for which $G_{xx}^\uparrow = 0$. In that case, $H_f(\omega)$ is directly measured. The amplitude and the phase of the antenna transfer functions are presented in Figure 3 as a function of both frequency and antenna height above the copper sheet. The regions of maximum gain of the antenna (i.e., corresponding to the minimum values of $|H_i|$ and maximum values of $|H_t|$) can be observed. For $h > 110$ cm, H_i tends to the measurement in free space conditions, which shows that the model is valid above that height. In addition, this indirectly means that the transfer functions H and H_f are also correctly determined. The progressive shifts in the transfer functions that are observed for the pattern of $|H_i|$ as well as for that of the phase of the three transfer functions for $h < 110$ cm indicate that the hypotheses behind the antenna model are not valid anymore (near field effects affecting the spatial distribution of the measured electric field). In that respect it is worth noting that we observed that the minimal height above which the antenna model holds corresponds to the size of the antenna aperture. For antennas operating at higher frequencies, whose size is proportionally smaller, this minimal height is smaller. Knowledge of the antenna transfer functions allows to filter the antenna effects out of the $S_{11}(\omega)$ measurements, thereby providing the measured Green's functions ($G_{xx}^{\uparrow meas}$), which can then be inverted to retrieve the soil electrical properties. It is worth noting that, theoretically, the antenna transfer functions should be determined only once, as they only depend on the antenna itself and the calibration plane for

the VNA.

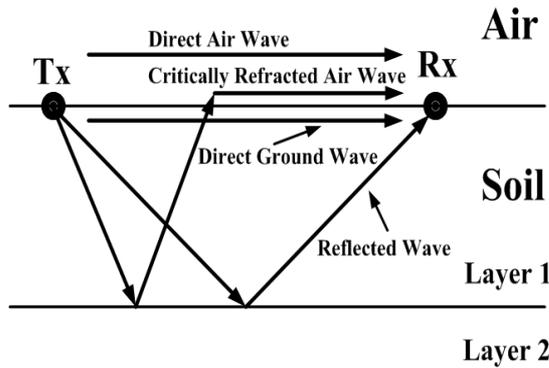


Figure 1 : Propagation paths of electromagnetic waves in a two layers soil.

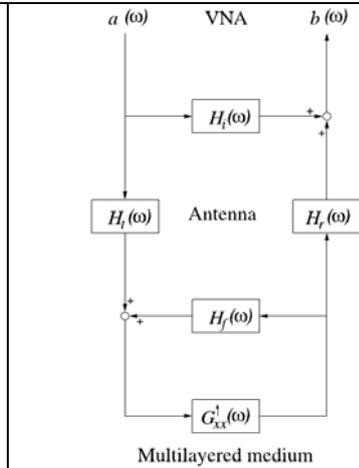


Figure 2 : Block diagram representing the VNA-antenna-multilayered medium.

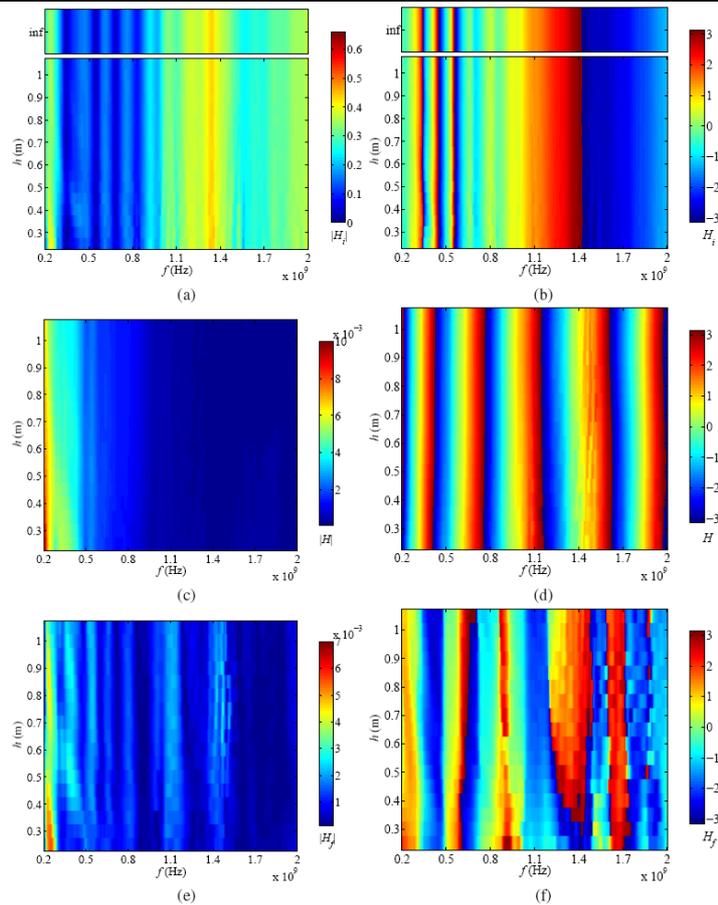
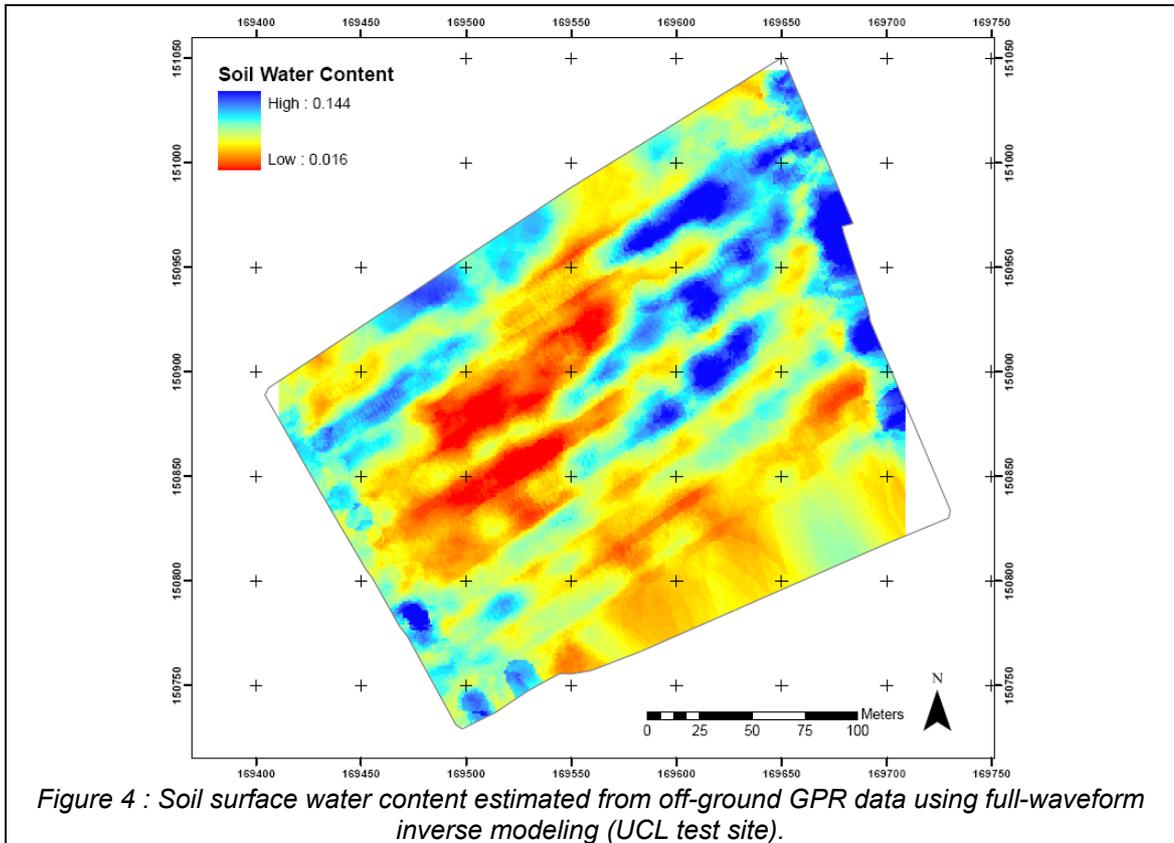


Figure 3 : Amplitude and phase of the antenna transfer functions H_i , $H = H_t H_r$, and H_f determined for different averaged elevations h of the antenna above a metal sheet. Measurement at 180 cm height was assumed to correspond to free-space conditions.

GROUND PENETRATING RADAR
<p>Retrieval of the soil properties</p> <p>For the traditional GPR methods, as explained above, the soil dielectric permittivity is estimated either from the Fresnel reflection coefficient or from the wave propagation velocity assuming straight-ray propagation. When the soil dielectric permittivity is known, it is correlated to the volumetric water content using a petrophysical model such as Topp's equation.</p> <p>For more advanced retrieval, e.g., with the GPR prototype, the constitutive parameters (ϵ_n, σ_n) and the thicknesses (h_n) of each subsurface layer are retrieved by inverse modelling, by minimizing an objective function defined as follows :</p> $\phi(\mathbf{b}) = \left \mathbf{G}_{xx}^{\uparrow meas} - \mathbf{G}_{xx}^{\uparrow mod} \right ^T \mathbf{C}^{-1} \left \mathbf{G}_{xx}^{\uparrow meas} - \mathbf{G}_{xx}^{\uparrow mod} \right \quad (8)$ <p>where $\mathbf{G}_{xx}^{\uparrow meas} = \mathbf{G}_{xx}^{\uparrow meas}(\omega)$ and $\mathbf{G}_{xx}^{\uparrow mod} = \mathbf{G}_{xx}^{\uparrow mod}(\omega, \mathbf{b})$ are, respectively, the measured and modelled Green's functions, $\mathbf{b} = [\epsilon_i, \sigma_i, h_i] (i=1, \dots, n)$, and \mathbf{C} is the measurement error covariance matrix. The minimization of the objective function is carried out using the global multilevel coordinate search (GMCS) algorithm combined sequentially with the classical Nelder-Mead simplex (NMS) algorithm. It is worth noting that parameter retrieval with inverse modelling will be successful only if enough information is contained in the data with respect to the parameters to be estimated, namely, the inverse problem should be well-posed. In case the inverse problem is not well-posed, it should be regularized using other sources of information such as a priori knowledge (e.g., from soil textural maps or expected parameter ranges) or data provided by other sensors. Typically, the inverse problem is well-posed for reconstructing up to two layered media.</p> <p>Figure 4 shows an example of application in real field conditions, where soil surface water content was mapped, by focusing the GPR full-waveform inversion on the surface reflection only in the time domain. In that case the parameters to be retrieved are the antenna height above the ground and the surface dielectric permittivity. Given the simplicity of the inverse problem, it is well-posed (the inverse solution is unique) and readily solved in less than one second on a single processor using local optimization.</p>



GROUND PENETRATING RADAR

Operational conditions and limits

Traditional GPR is operated by sliding the antennas in contact with the ground, the antennas being mounted on a sledge (as illustrated in Table 1). Measurements are automated and controlled by a laptop. Data are visualized in real time, as illustrated in Figure 5. The survey velocity is typically 5 km/h, but much larger velocities are possible, depending on the radar settings and soil conditions (smooth vs. rough). A survey wheel is usually used for accurate positioning, but the system can also be combined with a differential GPS. Figure 5 shows GPR data acquired with the SIR-20 with 400 MHz antennas from GSSI in the vineyards of Cheval Blanc (Saint-Emilion, France). Both monostatic (zero-offset) and bistatic (offset of about 80 cm between the transmitting and receiving antennas) data were simultaneously recorded. We can observe the presence of several soil layers, which appear at different propagation times along the transect. The reflection around 20 ns delineates the arable soil from the underlying layers. The strength of that reflection indicates a strong contrast between the two soil layers. Between about 30 and 60 meters along the transect, the reflection is weaker and more heterogeneous. Quantitative results from these data are not yet available.

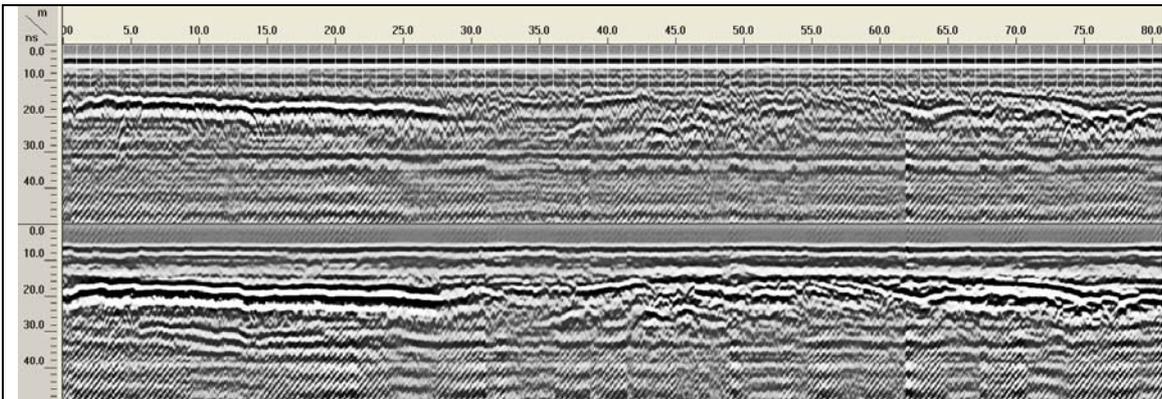


Figure 5 : GPR data acquired with a time-domain radar (SIR-20 with 400 MHz antennas from GSSI) in the vineyards of Cheval Blanc (Saint-Emilion, France). Both monostatic (top) and bistatic (bottom) data were simultaneously recorded.

The proposed GPR prototype is mounted on a quad and is fully computer controlled. Positioning is performed using an RTK differential GPS. EMI data can be acquired simultaneously. The antenna is situated at two meters behind the quad to avoid ambiguous reflections from the quad, and is situated at about 110 cm above the ground to ensure accurate modelling of the radar data. Measurements are acquired in real time, typically with a velocity of 5 km/h, depending on the VNA performances and its settings. Figure 5 shows an example of operational set up in the field where the radar data are simultaneously acquired with EMI data (from the EM38 in this case).



Figure 6 : GPR prototype mounted on quad for real-time mapping of the soil dielectric permittivity (Selhausen test site of FZJ).

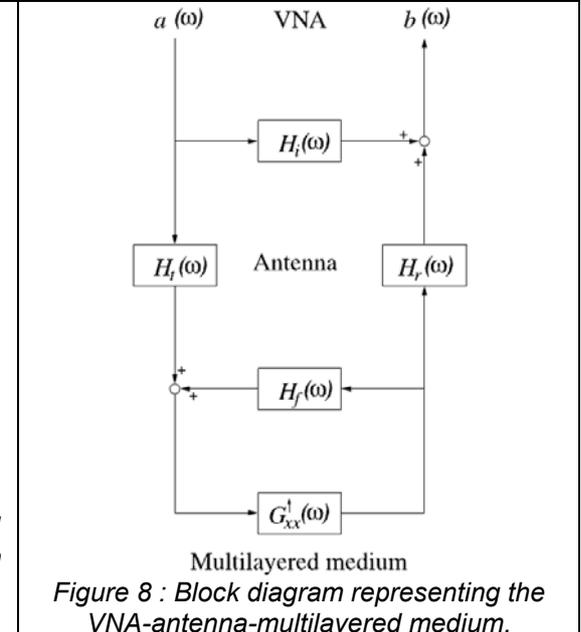
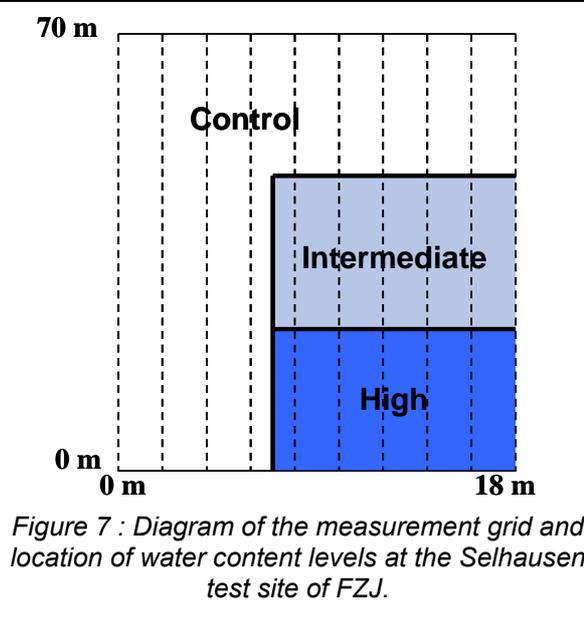
4. EM method

ELECTROMAGNETIC INDUCTION	
Sensor description	
<p>Three electromagnetic induction (EMI) sensors are investigated in DIGISOIL: (1) the EM38 from Geonics, that is commonly used for mapping the soil electrical conductivity in agricultural fields, (2) the new EMP-400 Profiler from GSSI, and (3) a prototype sensor that is being developed by FZJ and UCL within the DIGISOIL project. The main characteristics of these sensors are depicted in Table 1. A major advantage of the EMP-400 over the EM38 is its ability to simultaneously measure at three different operating frequencies, thereby allowing for different characterization scales and depths. In addition, these frequencies can be selected arbitrarily in the range 1-16 kHz. As this tool has been made available only recently, there are not yet scientific studies which demonstrated its performances compared to the traditional EM38. Compared to the two commercial systems, the searched advantages with the prototype development are (1) a more robust calibration procedure to ensure repeatability of the measurements, and (2) a fully mechanistic signal processing based on three-dimensional (3-D) electromagnetic forward modelling and inversion.</p>	
	<p>EM38 (Géonics) Dimensions: $L=1.06\text{ m} \times l=0.036\text{ m} \times h=0.15\text{ m}$ Operating frequency: 14.6 kHz Operating modes⁽¹⁾: H-H, V-V Conductivity range: $0\text{-}1\ 000\text{ mS}\cdot\text{m}^{-1}$ Tx-Rx spacing: 1.0 m</p>
	<p>EMP-400 Profiler (GSSI) Dimensions: $L=1.46\text{ m} \times l=0.124\text{ m} \times h=0.24\text{ m}$ Operating frequencies: from 1 to 16 kHz, up to 3 frequencies simultaneously Operating modes⁽¹⁾: H-H, V-V Conductivity range: $0\text{-}1\ 000\text{ mS}\cdot\text{m}^{-1}$ Tx-Rx spacing: 1.217 m</p>
	<p>EMI prototype (FZJ and UCL) Dimensions: diameter=$0.33\text{ m} \times h=0.24\text{ m}$ Operating frequencies: from 1 to 50 kHz Operating modes⁽¹⁾: V-V Conductivity range: $0\text{-}1\ 000\text{ mS}\cdot\text{m}^{-1}$ Tx-Rx spacing: 0 m (zero-offset)</p>
<p><i>Table 1 : Description of the three EMI sensors that are investigated in DIGISOIL. (1)H-H: transmitter and receiver horizontal dipole orientation; V-V: transmitter and receiver vertical dipole orientation.</i></p>	

ELECTROMAGNETIC INDUCTION	
Signal processing	
EM38 and EMP-400 Profiler	
<p>After appropriate calibration of the instrument, EM38 and EMP-400 Profiler provide directly values of the soil apparent electrical conductivity (i.e., depth-weighted average of soil electrical conductivity) for each mode of operation (i.e., H-H or V-V) and, for the Profiler, for each selected frequency. Yet, depending on the operating conditions, corrections of the sensor output values may be required for proper soil electrical conductivity measurements. First, values of electrical conductivity provided by the instruments rely on the low induction number (β) assumption (i.e., $\beta \ll 0.5$), with :</p>	
$\beta = \frac{s}{\delta} \quad (9)$	
<p>where s (m) is the Tx-Rx spacing, and δ (m) is the skin depth, namely the soil depth at which the amplitude of the primary magnetic field has been attenuated to $1/\exp$ the amplitude at the surface. The skin depth is given by :</p>	
$\delta = \sqrt{\frac{2}{\omega\mu_0\sigma}} \quad (10)$	
<p>where ω (rad) is the angular frequency, $\mu_0 = 4\pi \times 10^{-7}$ (Hm⁻¹) is the magnetic permeability of free-space, and σ (Sm⁻¹) is the soil electrical conductivity. Under low induction number conditions, the soil apparent electrical conductivity may be expressed as (McNeill, 1980) :</p>	
$\sigma_a = \frac{4}{\omega\mu_0 s^2} \left(\frac{H_s}{H_p} \right)_Q \quad (11)$	
<p>where $(H_s/H_p)_Q$ (dimensionless) is the ratio of the quadrature components of the secondary and the primary magnetic fields at the receiver coil. This expression corresponds to the electrical conductivity indicated by the instruments. However, beyond the low induction zone, this asymptotic linear relationship does not hold anymore and a correction has to be applied, such as that proposed by van der Kruck et al. (2000) using the information contained in the ratio of the in-phase components of the secondary and primary magnetic fields at the receiver coil, $(H_s/H_p)_I$ (dimensionless), also measured and recorded by the sensors. A second correction to be applied relates to the variation of soil electrical conductivity as a function of soil temperature (Slavich and Petterson, 1990). As a result, measurements are standardized according to a reference temperature, usually 25°C (Sheets and Hendrickx, 1995) :</p>	
$\sigma_{25} = \sigma \left(0.447 + 1.4034e^{-(T/26.815)} \right) \quad (12)$	
<p>where σ_{25} (Sm⁻¹) is the standardized σ at a temperature of 25°C and T (°C) is the soil temperature.</p>	
<p>As an example of EM38 and EMP-400 Profiler data processing, Figure 9 presents electrical conductivity measurements performed at the Selhausen test site (Germany). Both vertical and horizontal dipole modes were considered for each instrument, and three frequencies were selected for the Profiler (5 kHz, 10 kHz, and 15 kHz). Measurements were carried out on a grid, every four meters along ten transects separated each by two meters. Some parts of the study site were irrigated prior to the measurements, thereby providing three water content distinct ranges (Figure 1): 'control' (natural water content), 'intermediate' (natural + 5 mm irrigation), and 'high' (natural + 10 mm irrigation). The low induction number condition being satisfied ($\beta \approx 0.02$), data processing was carried out without considering the in-phase component. Yet, soil</p>	

temperature during the measurements averaged 19.8°C and electrical conductivity data were corrected using Equation (12), leading to higher corrected data compared with the raw data. In each case, gradients of soil electrical conductivity were observed over the study area, with a generally rather good correspondence with the water content levels. For each operating mode, results for the EM38 are quite similar to those of the Profiler at 15 kHz and 10 kHz, while electrical conductivities from Profiler at 5 kHz tend to be higher and more spatially variable. The investigated depth tending to increase as the frequency decreases (see Equation (10)), higher electrical conductivity values at 5 kHz compared with the other frequencies would reflect a trend of soil electrical conductivity to increase with depth; the lower spatial correlation at 5 kHz could at least partly result from a greater influence of the environment (other sensors and probes installed at the study site) at that frequency compared with the others. The presence of more conductive layers at lower depth compared to the surface is confirmed by higher electrical conductivities obtained from vertical dipoles than from horizontal dipoles, the former operating mode being more sensitive to lower depths than the second one.

EMI prototype

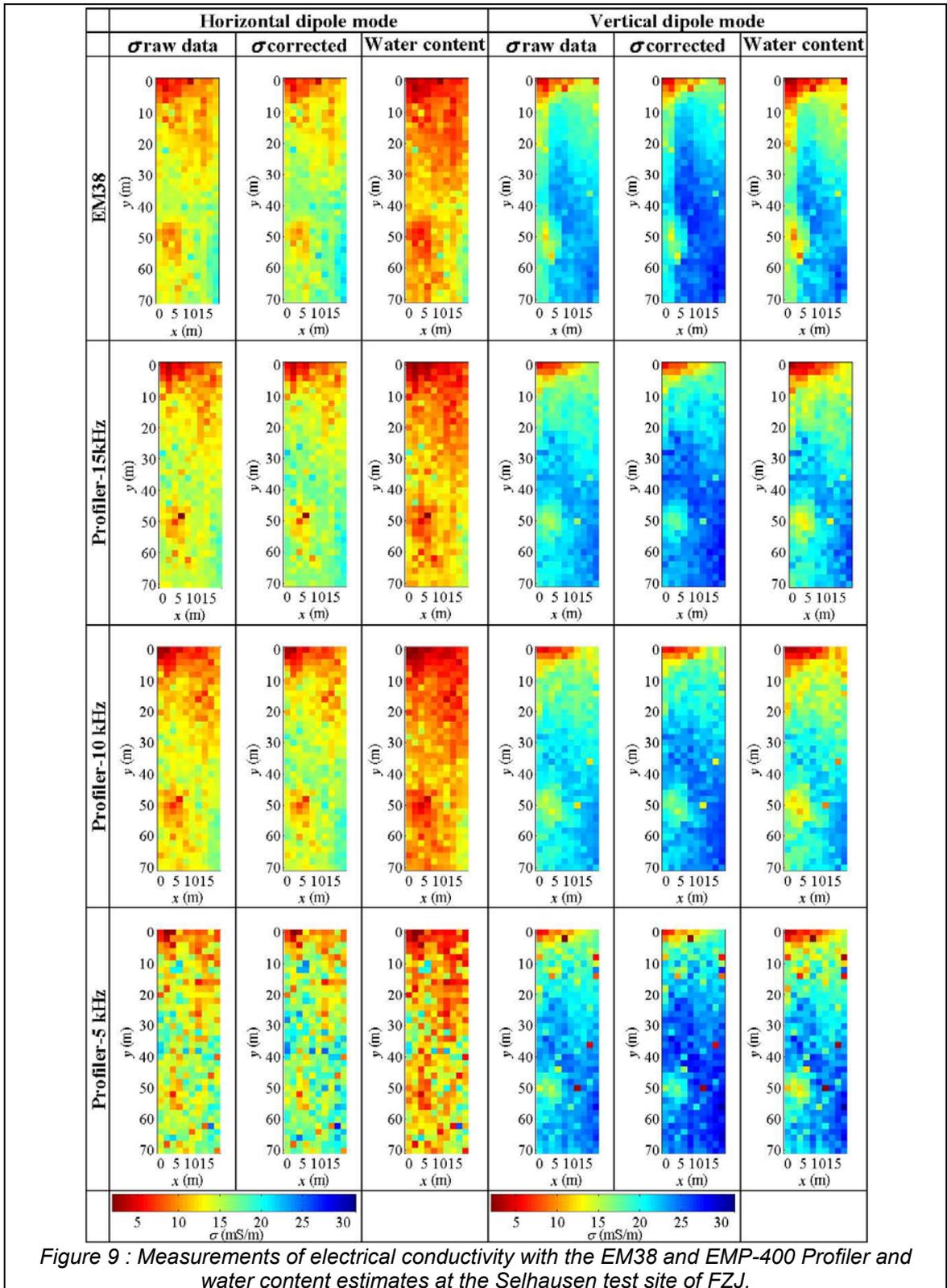


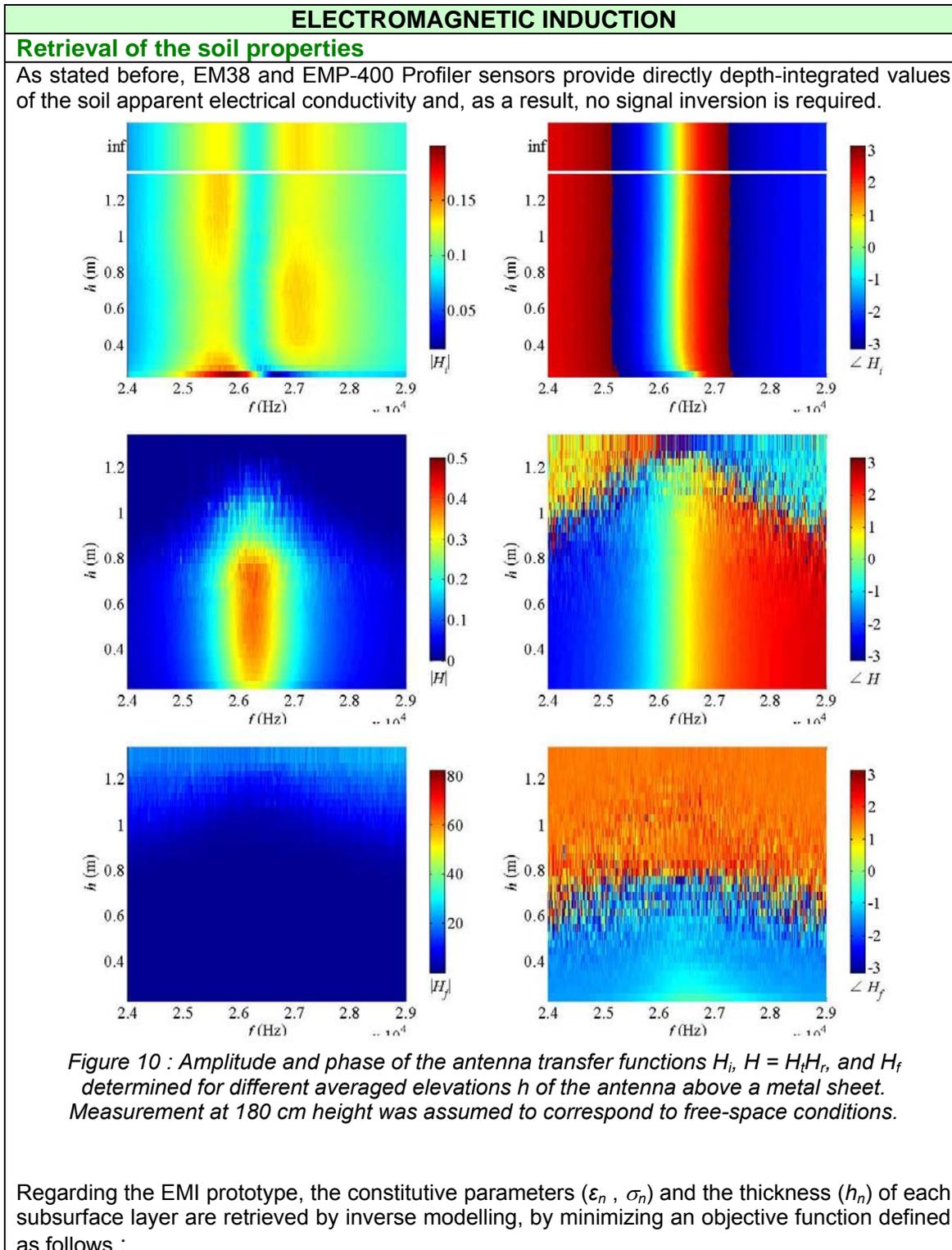
For the EMI prototype, the measured quantity is the frequency-dependent complex ratio, $S_{21}(\omega)$, between the received signal at port 2 of the vector network analyser (VNA) and the transmitted signal at port 1. This ratio contains both the subsurface response and the antenna effect information. It may be represented by the block diagram shown in Figure 2 and accordingly expressed in the frequency domain by the following equation (Lambot *et al.*, 2004; Moghadas *et al.*, submitted):

$$S_{21}(\omega) = \frac{b(\omega)}{a(\omega)} = H_i(\omega) + \frac{H_i(\omega)G_{zz}^{\dagger}(\omega)H_r(\omega)}{1 - H_f(\omega)G_{zz}^{\dagger}(\omega)} \quad (13)$$

where $b(\omega)$ and $a(\omega)$ are, respectively, the backscattered and incident waves at the VNA reference calibration planes; $H_f(\omega)$, $H_i(\omega)$, $H_r(\omega)$ and $H_t(\omega)$ are the characteristic antenna transfer functions accounting for the antenna effects and antenna-soil interactions; G_{zz}^{\dagger} is the transfer Green's function of the air-subsurface system modelled as a 3-D multilayered medium, each n th layer being considered as homogeneous and characterized by its dielectric permittivity (ϵ_n), electrical conductivity (σ_n) and thickness (h_n). The magnetic permeability is assumed to be constant and equal to the permeability of free space. The characteristic antenna transfer

functions can be determined by solving a system of equations as (7) for different model configurations. To this end, we use well defined model configurations, i.e., with the antenna situated at k different heights above a perfect electrical conductor (copper sheet), so that the Green's functions G_{zz}^{\uparrow} can readily be computed and the corresponding $S_{21}(\omega)$ can be measured in a standard way. The system of equations should be overdetermined ($k > 3$) to ensure a well-defined and accurate solution (the equations may not be fully independent for the whole frequency range, depending on the measurement heights). Function $H_i(\omega)$ can also be determined in an independent way, by performing measurements in free space conditions for which $G_{zz}^{\uparrow} = 0$. In that case, $H_i(\omega)$ is directly measured. The amplitude and the phase of the loop antenna transfer functions are presented in Figure 3 as a function of both frequency and antenna height above the copper sheet. The region of maximum gain of the antenna (i.e., corresponding to the minimum value of $|H_i|$ and maximum value of $|H|$) is at 26.3 kHz, which corresponds to the resonant frequency of the system. For $h > 30$ cm, H_i tends to the measurement in free space conditions, which shows that the model is valid above that height. In addition, this indirectly means that the transfer functions H and H_f are also correctly determined. The shifts in the resonant frequency towards higher frequencies observed for the pattern of $|H_i|$ as well as for that of the phase of the three transfer functions for $h < 30$ cm indicate that the hypotheses behind the antenna model are not valid anymore (near field effects affecting the spatial distribution of the measured magnetic field). Apart from the noise, the highest values of $|H|$ are observed between about 30 cm and 80 cm height. Moreover, $|H|$ is relatively constant within this height interval, while rather strong variations are observed for higher antenna elevations. As $|H|$ is theoretically constant (does not depend on the calibration height), this is to be attributed to a progressive decrease in the signal-to-noise ratio. This effect is also particularly well observed in the phase of H and H_f , for which noise increases with both calibration heights and distance from the antenna resonant frequency. These observations reveal that model validity and antenna gain are acceptable between 30 and 80 cm height for this particular EMI setup. Knowledge of the antenna transfer functions allow to filter the antenna effects out of the $S_{21}(\omega)$ measurements, thereby providing the measured Green's functions ($G_{zz}^{\uparrow meas}$), which can then be inverted to retrieve the soil electrical properties. It is worth noting that, theoretically, the loop antenna transfer functions should be determined only once, as they only depend on the antenna itself and the calibration plane for the VNA.





$$\phi(\mathbf{b}) = \left| \mathbf{G}_{zz}^{\uparrow meas} - \mathbf{G}_{zz}^{\uparrow mod} \right|^T \mathbf{C}^{-1} \left| \mathbf{G}_{zz}^{\uparrow meas} - \mathbf{G}_{zz}^{\uparrow mod} \right| \quad (14)$$

where $\mathbf{G}_{zz}^{\uparrow meas} = \mathbf{G}_{zz}^{\uparrow meas}(\omega)$ and $\mathbf{G}_{zz}^{\uparrow mod} = \mathbf{G}_{zz}^{\uparrow mod}(\omega, \mathbf{b})$ are, respectively, the measured and modelled Green's functions, $\mathbf{b} = [\epsilon_i, \sigma_i, h_i]$ ($i=1, \dots, n$), and \mathbf{C} is the measurement error covariance matrix. The minimization of the objective function is carried out using the global multilevel coordinate search (GMCS) algorithm combined sequentially with the classical Nelder-Mead simplex (NMS) algorithm (Lambot *et al.*, 2002). As an example of current results obtained with this new EMI procedure, the antenna heights above the copper sheet were inversely estimated from corresponding measurements performed for the characterization of the antenna transfer functions (see preceding section). The estimated heights are compared with the measured values in Figure 4. The four red triangle symbols represent the heights used for the antenna calibration, i.e., to determine the antenna transfer functions. As clearly seen, a very good agreement is observed between the inversely estimated and directly measured heights. Indeed, antenna heights could be retrieved with sub-millimetre accuracy within the optimal efficiency 0.30 m – 1.00 m interval, while large discrepancies were observed for lower and higher heights as a result of model inadequacy (the model hypotheses do not hold anymore) and low signal-to-noise ratio, respectively. These results are quite promising with respect to the potential of this EMI prototype for soil characterization. Nevertheless, preliminary laboratory experiments on water with different salinity levels revealed that the sensitivity of this system still needs to be improved for accurate measurements in the range of electrical conductivities encountered in soils.

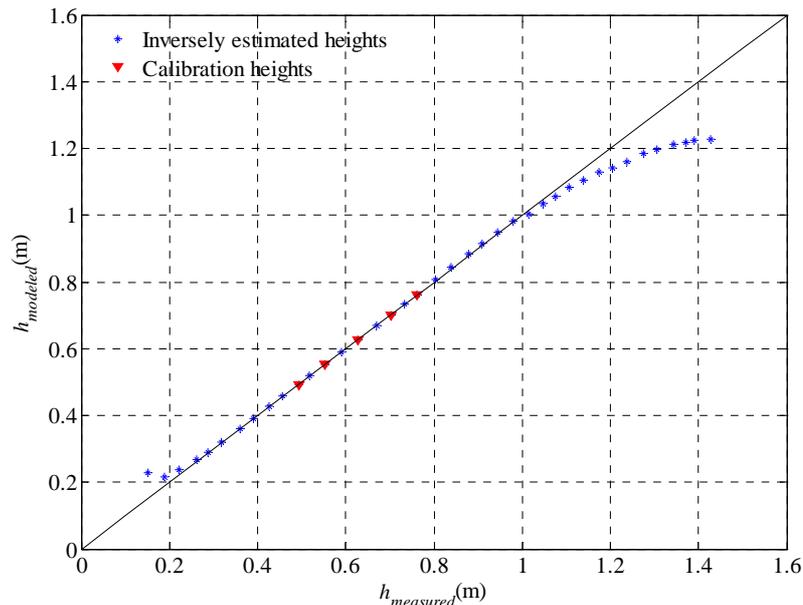


Figure 11 : Comparison of inversely estimated heights (h_{modeled}) and directly measured heights (h_{measured}). The heights used for the antenna calibration are represented by the red triangles.

Furthermore, once proper estimates of the soil electrical conductivity are obtained, they can be used to assess physico-chemical soil properties using petrophysical relationships. Indeed, soil electrical conductivity is influenced by soil water content, clay content, bulk density and chemical composition of the soil solution. Several models are proposed in the literature to relate soil electrical conductivity to soil physico-chemical properties. Rhoades *et al.* (1976) presented a simple linear model relating soil electrical conductivity to liquid-phase electrical conductivity, to water content and to surface conductivity :

$$\sigma = (a\theta^2 + b\theta)\sigma_w + \sigma_s \quad (15)$$

where σ (Sm^{-1}) is the bulk soil electrical conductivity, θ (m^3m^{-3}) is the soil water content, σ_w (Sm^{-1}) is the soil solution electrical conductivity, σ_s (Sm^{-1}) is the electrical conductivity of dry soil, and a and b are soil specific empirical parameters. An other model is the Archie's law (1942) :

$$\frac{\sigma}{\sigma_w} = \phi^m S^n \quad (16)$$

where, as previously, σ (Sm^{-1}) and σ_w (Sm^{-1}) are the soil and soil solution electrical conductivity, respectively, and ϕ represents the soil porosity, S is the degree of saturation, m is the coefficient of tortuosity, and n is a fitting parameter. Since then, variations of the Archie's law have been developed (Erwing and Hunt, 2006). Shah and Singh (2005) notably proposed a modified formulation :

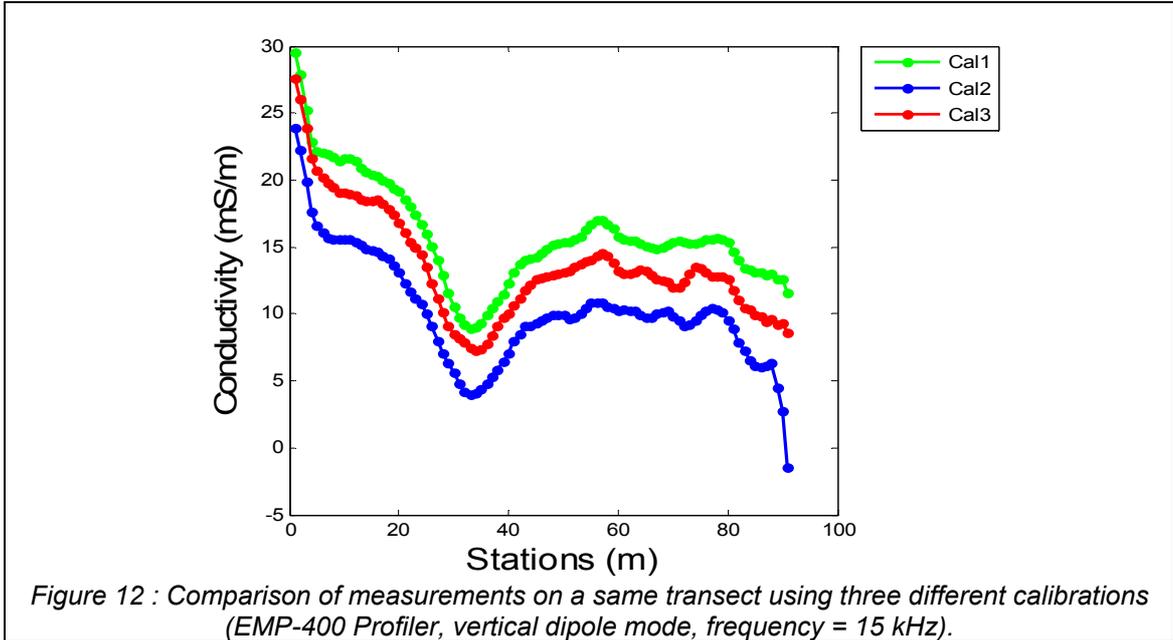
$$\frac{\sigma}{\sigma_w} = c\theta^m \quad (17)$$

where c and m are fitting parameters depending primarily on the particle size characteristics of the soil. The same authors proposed the following empirical relationships between c and m , and the clay content of a soil :

$$c = i \cdot \text{Clay}^j \quad (18)$$

$$m = k \cdot \text{Clay}^l \quad (19)$$

where Clay is expressed as percent by volume. For soils with Clay > 5 %, they defined values of $i = 0.6$, $j = 0.55$, $k = 0.92$ and $l = 0.2$. If clay content is less than 5 %, c and m are constant and equal to 1.45 and 1.25, respectively. As an example, water content was estimated from soil electrical conductivity measurements carried out at the Selhausen test site (see above) using the modified formulation of the Archie's law. The percentage of clay was rather stable all over the study area and the average value of 15% was used in equations (18) and (19) and the soil solution electrical conductivity was considered as constant. As expected from the electrical conductivity values, spatial variations of water content are observed with respect to the irrigation levels. Water content values of both operating modes are particularly contrasted for the 'intermediate' and 'high' irrigation levels, with higher values of water content for vertical dipoles compared with horizontal dipoles (i.e., at lower depth compared with the surface), which would arise from the infiltration of the irrigated water along the vertical soil profile. Nevertheless, applying these empirical relationships to our soil apparent electrical conductivity only provides rough estimations of the soil water content, integrated all over the depths of sensitivity of the sensors (depending on the Tx-Rx spacing, on the operating mode and on the frequency). Moreover, accurate estimates of water content would require detailed information on the horizontal and vertical spatial variations of the other physico-chemical soil properties affecting soil electrical conductivity. Another limitation in using commercial sensors as EM38 and EMP-400 Profiler is the lack of robustness and repeatability of their calibration. As an illustration, Figure 12 presents the comparison of measurements performed along a same transect in identical conditions (i.e., with 3-minute interval) after three different calibrations of the EMP-400 Profiler, operating in vertical dipole mode for 15 kHz frequency. These results clearly show a weak repeatability of the measurements amongst the three calibrations, while the general trend of the conductivity values along the transect is quite similar. This underlines a major advantage in using VNA technology for the EMI prototype we are developing, as VNA calibration is robust and defined as an international standard.



ELECTROMAGNETIC INDUCTION

Operational conditions and limits

The proposed EMI prototype has not yet been tested in field conditions, but should operate with the loop antenna situated at about 30-40 cm above the soil surface. The EM38 and EMP-400 Profiler can operate with the sensors on the ground surface or with some height above the ground (typically 10-20 cm for real-time mapping). The measurement height will influence the characterization depth. Figure 5 shows an example with the EM38 mounted in front of a quad. The EM38 data are automatically acquired in real-time with a laptop and data are simultaneously geo-referenced with a differential GPS.



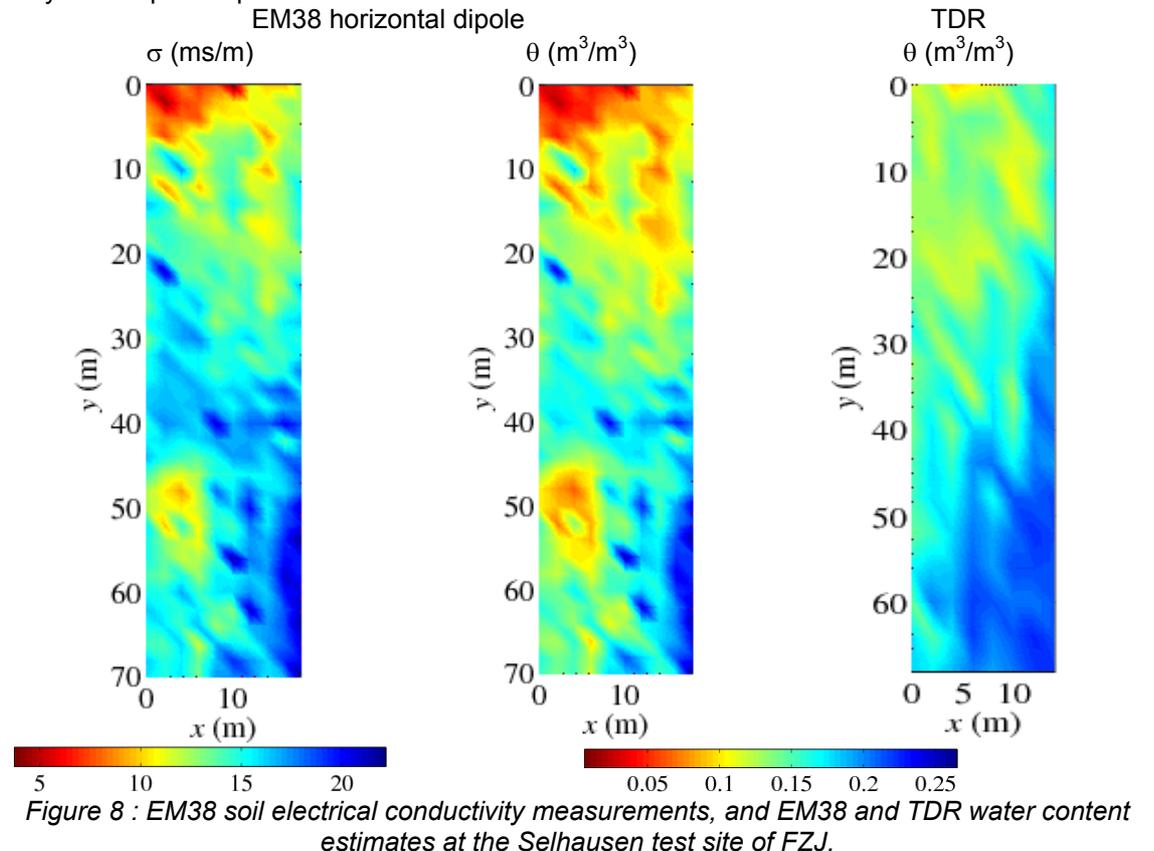
Figure 13 : EM38 mounted in front of a quad for real-time mapping of the soil electrical conductivity (Selhausen test site of FZJ). A GPR system is mounted at the back of the quad. A computer automatically controls the measurements.

If these are not taken into account in the sensor calibration, belowground and/or aboveground

metallic objects in the direct vicinity ($\approx 1-2$ m) of the instrument may influence EMI soil electrical conductivity measurements. In addition, the magnetic field generated by neighbouring high voltage power lines may also significantly affect EMI measurements. Ideally, these potential sources of influence should be removed prior performing the measurements, or great care is requested in the interpretation of the corresponding data.

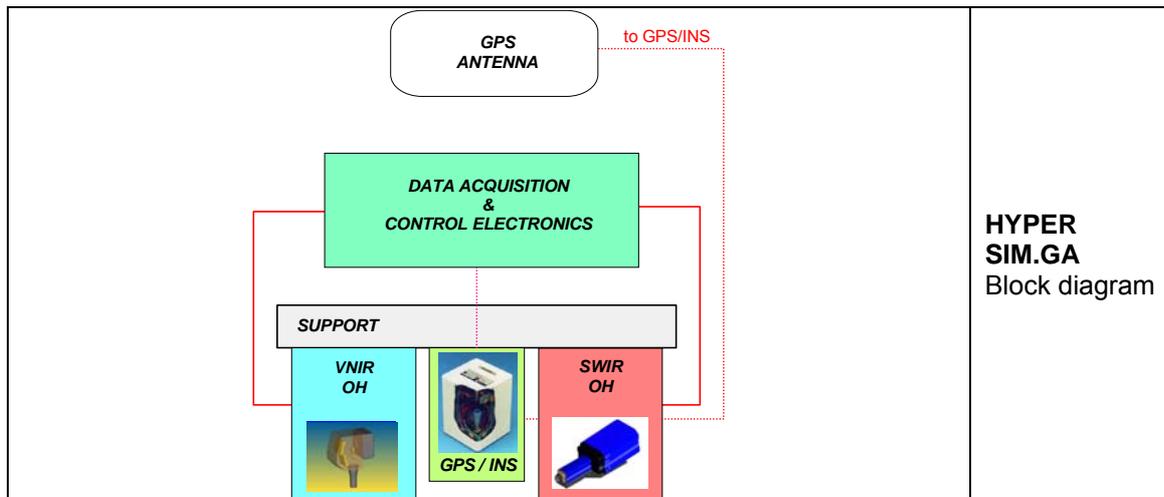
Field testing

As an example of field experimental data, Figure 8 presents horizontal dipole EM38 electrical conductivity measurements with their corresponding water content estimations (using equations (10) and (11)) for the Selhausen irrigation experiment (see above), and compare them to estimates of water content from TDR measurements performed simultaneously. For both techniques, similar gradients of water content are observed over the study area. Nevertheless, the range of water content values is slightly larger for EM38 results compared with TDR. The differences between both estimations are relatively larger in the ‘control’ zone (see Figure 2), especially at the upper part of the site and in an area at around 20 m from the lower side. For this latter area, the proximity of a meteorological station with a metallic mast and several sensors might have influenced the EMI measurements, which may at least partly explain the differences. The observed differences may also arise from the difference of investigation depth of both techniques, with a significant sensitivity of EM38 measurements in horizontal dipole mode to electrical conductivity of soil layers up to more than 70 cm depth (McNeill, 1980), while TDR measurements integrate the first 10 cm of the soil profile. Local variations of the soil clay content may also affect the EMI measurements and the corresponding water content estimations; however, the clay content is relatively stable with variations lower than 1% all over the study area. Finally, the lack of robustness of the calibration of the instrument (see above) may also explain a part of the differences between EMI and TDR water content estimations.



5. Hyperspectral method

HYPER SPECTRAL	
Sensor description	
<p>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 a the UNIFI ultra-light aircraft FOLDER at about 1000 m of altitude. The hyperspectral HYPER SIM.GA is composed of two optical heads (Figure 3.80):</p> <ol style="list-style-type: none"> 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 <p>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.</p> <p>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.</p> <p>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.</p> <p>The HYPER was first flows by Galileo Avionica on December 2005 for a demonstration campaign including ground surveying and was tested on 2006-2007 during ESA-AirFire and ESA-CEFLES2 campaigns.</p>	
	
<p>HYPER SIM.GA UNIFI FOLDER installation and bottom view of optics</p>	



HYPER SIM.GA
Block diagram

	VNIR Spectrometer	SWIR Spectrometer
Spectral Range	400-1000nm	1000 –2450nm
Spectral Sampling	1.2nm	5.8nm
Spectral bands	512	256
Spatial pixels	1024	320
IFOV per pixel	0.7mrad	1.33mrad
FOV	±20.71°	±12.04°
GSD@H=1000m	0.7m	1.33m
SWATH@H=1000m	722m	425m
F#	2.3	2.3
Digital resolution	12 bit	14 bit
Sensor	Frame transfer CCD	CMT cooled @200K
Max Operating Frame Rate (programmable)	57Hz	100 Hz
DIGISOIL Operating Frame Rate	42 Hz	22 Hz
Max Data Rate	57MB/s	5MB/s
Operating Data Rate	44MB/s	
Storage Capacity	200GB	
Autonomy	1h 15'	

SIM.GA HYPER Main characteristics

Table 1 : Description of the hyperspectral sensor and airborne installation investigated in DIGISOIL. From top to bottom: SIMGA system, block diagram, main characteristics

HYPER data processing

SIM-GA HYPER flight lines data are acquired during DIGISOIL campaigns over agricultural hilly areas in Tuscany (Italy) according to the agreed flight plan. All imagery are 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 are 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 measurements. 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)

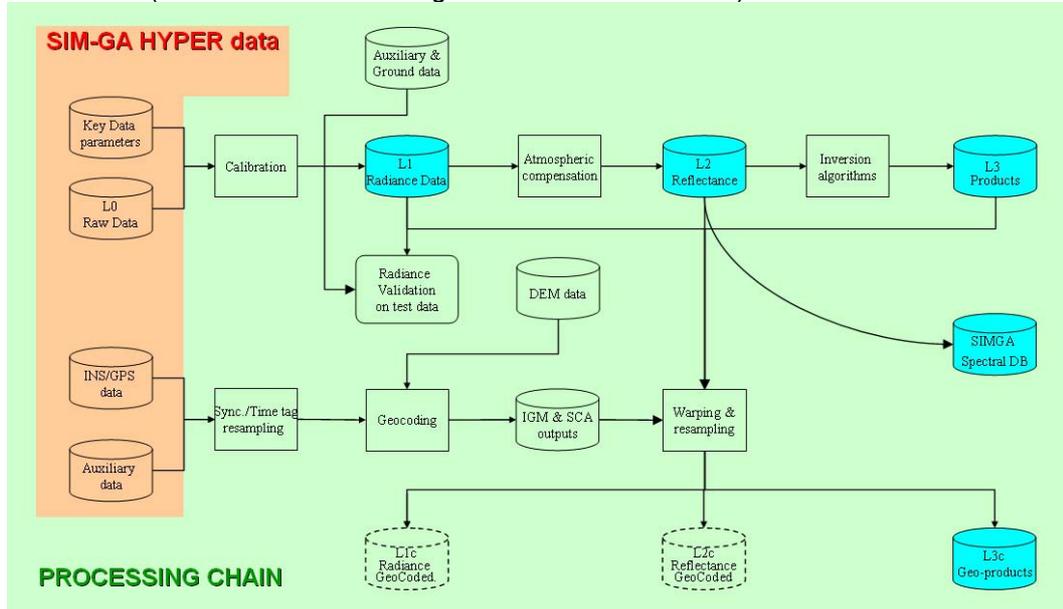


Figure 1: the hyperspectral processing chain

The HYPER data processing chain is outlined in Figure 1, where the flow diagram from the pre-processor, through the Level 1 processor, the auxiliary ground data up to intermediate (Radiance and geocoding outputs) is described. The processing of HYPER final products (Level 2-3 products) is performed by UNIFI and UCL in the framework of DIGISOIL team.



Figure 2: pre-processor main window

The pre-processor tasks deal with the consistency and quality/error check on the collected data set. In Figure is shown the pre-processor main window for the execution of the available check list.

Level 1 Processor description

The L1 processor task deals with the conversion of raw DN data into radiance physical units applying key data parameters from calibration and characterisation laboratory measurements. The Level 1 Processor is an IDL based project where inputs and outputs are shown in Table.

	INPUTS	OUTPUTS
RAW DATA	IMAGE .raw data (VNIR / SWIR) Header file IMAGE .cam file IMAGE .raw DARK data (VNIR / SWIR) DARK header file DARK.cam file	Radiance calibrated Image Header ENVI file Procedure LOG.html
KEY DATA PARAMETERS	Defective Pixels Map Flat Field coefficients Matrix Gain Coefficients Array Central Wavelength Array FWHM Array	

First, the raw data are corrected for the “exposure time” and the “dark signal”.

• Exposure time correction

The exposure time is in-flight programmable in order to avoiding saturation as well as under-exposition (low dynamic and signal). Usually it is set at a default value of 20 ms for VNIR and 10 ms for SWIR. The processor performs the exposure time normalisation as follows:

$$(DN_{i,j})_{@1s} = \frac{DN_{i,j}}{ExpTtms}$$

• Dark correction

During flight campaign, dark images are recorded at the beginning and at the end of each acquisition sequence. The Processor automatically detects these dark images searching for the temporal nearest dark file. In the case of SWIR images, the nearest temperature dark criterion yields the best correction of dark effects on the images (the sensor temperature during dark and image acquisition is recorded in the housekeepings data). A temporal dark average is applied in order to reduce the random noise:

$$(DN_{i,j})_{DARK SUB} = (DN_{i,j})_{@1s} - \langle (DN_{i,j}^{DARK})_{@1s} \rangle_{TEMP}$$

In the experienced flight conditions the dark signal was about 55 DN over a dynamic of 4096 (VNIR) e about 5400 DN over a dynamic of 16383.

• Defective pixel correction

Every image has to be masked with detected Defective Pixel Matrix. A matrix which contains the list and spatial/spectral position of defective pixels identified during calibration laboratory measurements. Level 1 Processor is able to point to this matrix for masking or restoring defective pixels along the processing chain according to the chosen criteria among the following:

- No correction
- Set to NaN
- Spatial interpolation (default)
- Spectral interpolation

• Flat field correction

Flat Field Coefficients matrixes are defined by laboratory calibration procedures. In the case of VNIR sensor two separately Flat Field are defined taking into account a) the two sensor halves are read by two different readout amplifiers and b) flat field matrix for low light levels where some non linearity of sensor response occurs. In this way the coherent noise striping is quite well reduced:

$$(DN_{i,j})_{FF\ CORR} = \frac{(DN_{i,j})_{DARK\ SUB}}{FF_{i,j}}$$

- Gain coefficient correction

Gain coefficients in W/(m² sr nm) are calculated during the calibration laboratory measurements. The coefficients are applied according to the formula:

$$(DN_{i,j})_{GAIN\ CORR} = (DN_{i,j})_{FF\ CORR} * GAIN_i$$

Geocoding processing description

The two main objectives of the geocoding processing are:

- Data correction and compensation for flight attitude parameters
- Data geolocation on the basis of DEM and geometrical sensor model

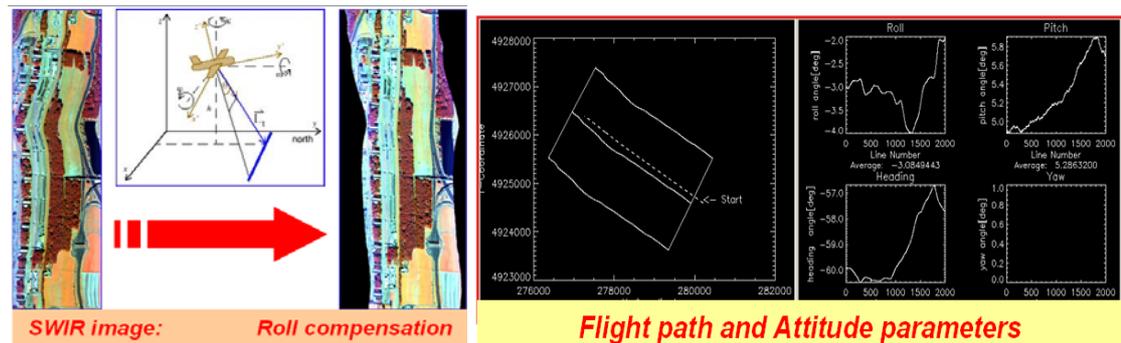


Figure 3: geocoding protocols

The geocoding processing (Figure 3) performs the following main tasks:

- Pre-processing of INS/GPS data

An IDL project was developed aiming to the conversion of binary navigation file from C-MIGITS III INS/GPS unit into text files, testing for files integrity and reporting possible errors. Then reference times for each HYPER frame are read from bytes where housekeeping data are stored and saved in two txt file where time samples are in the same number of HYPER frames for each VNIR and SWIR image block.

- Synchronization between INS/GPS and HYPER cameras

The task of HYPER and GPS/INS data synchronization is achieved through a dedicated electronic board ("Synch unit") which manages directly the GPS/INS and keeps the common time tag used to link VNIR and SWIR frames with the GPS/INS data. The first time tag received by each camera defines the start of the acquisition. Nonetheless an offset due to Synch Unit seems to be present between the time tag referred to frame grabber events (image frame acquisition) and GPS/INS data messages time tag (about 1 sec both for VNIR e SWIR). The file generated by the Synch Unit (data_log) is then processed and text files containing GPS, attitude and time data for each VNIR and SWIR image frames are saved.

- Boresight calibration of VNIR and SWIR cameras

The analysis of offsets results from different flight-path leads to the final averaged offset values. The same calibration procedure is performed for both VNIR and SWIR boresights. The task is performed on the basis of PARGE and an IDL procedure integrated in the geo-processor. The overall procedure aims to the geodetic refinement of INS/GPS data (datum projection, geoid undulation) and their compensation for boresight offsets before starting PARGE geocoding processing. Moreover, the UTM zone (number and letter) and the heading convergence because of possible offsets between HYPER and grid convergence are estimated:

$$\text{Headconv} = 180/\pi(\text{TrueHeading} - \text{atan}(\sin(\text{Lati}) * \tan(\text{Longi} - (\pi/180.0 * (\text{S}))))))$$

where S depends from UTM zone (e.g. S=3 for 31zone and S=-3 for 30 zone)

At the end of the procedure the boresight calibration is applied and attitude (roll, pitch, true head.) and GPS (Lat/Long/Alt) parameters are modified on the basis of datum, geoid undulation, heading convergence and boresight offsets. The considered inputs/outputs of the geocoding processing are reported in the following table:

INPUT	<ol style="list-style-type: none"> 1. HYPER VNIR or SWIR bsq raw data to be georeferenced 2. boresight corrected navigation data 3. VNIR/SWIR synchronised time tag data 4. VNIR/SWIR pixel size @ flight height 5. HYPER sensor model: <ol style="list-style-type: none"> a. FOV: 24.942 (SWIR) or 40.6 (VNIR) b. IFOV across and along track: 1.329 (SWIR) or 0.692 (VNIR) 6. DEM
OUTPUT	<ol style="list-style-type: none"> 1. IGM (Internal Geometry Map) 2. SCA (SCan Angle map) 3. GEO (georeferenced image cube) (Optionally)

The geocoding flow diagram is shown in Figure 4

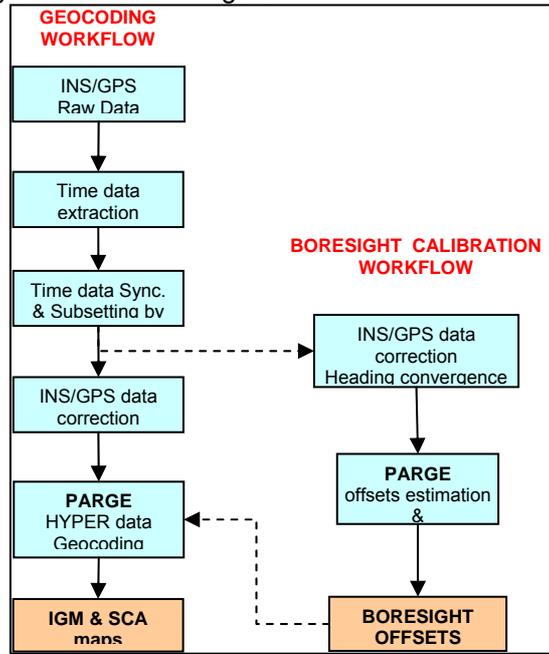
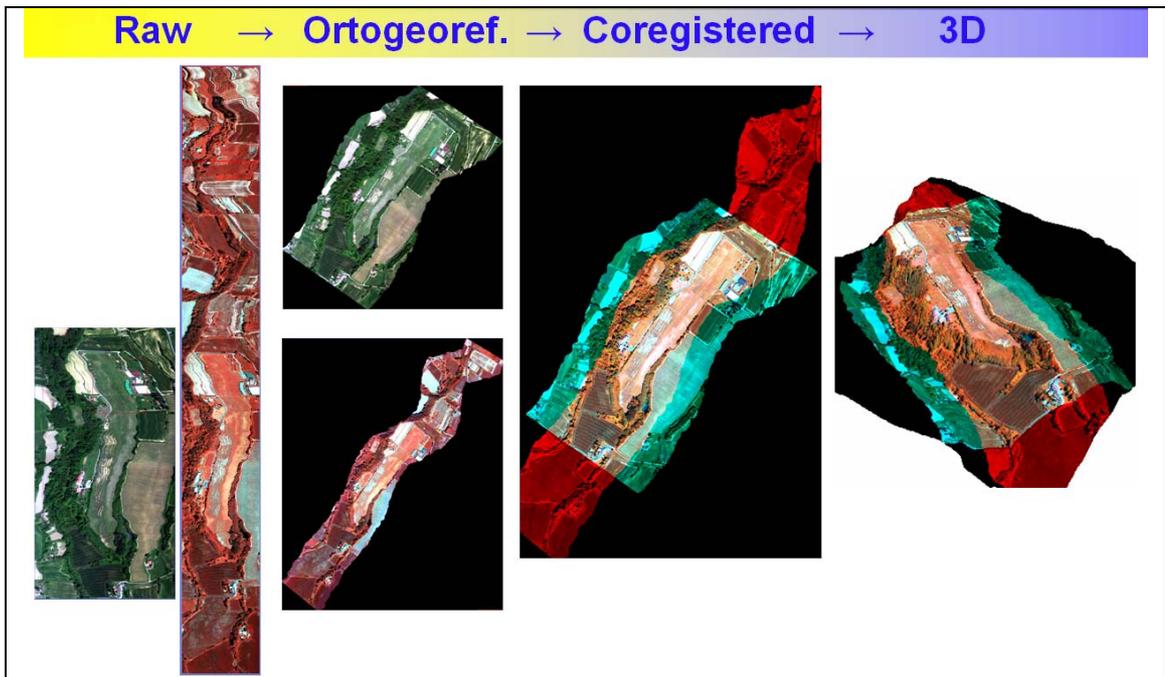


Figure 4: geocoding flow diagram

The effects of the geocoding processing on a sample dataset is depicted as RGB images in the following Figure, where raw data were compensated for flight attitude and georeferred, co-registered and then shown in 3D mode overlaid on the DEM.



HYPERSPECTRAL

Retrieval of the soil properties

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 et 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:

$$\begin{aligned} X &= TP^t + E \\ Y &= UQ^t + F \end{aligned} \quad (1)$$

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 indices. Such spectral indices have been developed e.g. to retrieve soil moisture (Whiting et al., 2004) and SOC content (Bartholomeus et al., 2008). They can be easily related to the biochemical constituents of the soil for which absorption feature is known and may show a greater robustness. These features are due to overtones and combinations of the fundamental vibrations of bonds C–O, C–H, N–H and O–H occurring in the Mid InfraRed (MIR). This strategy is particularly suitable when the property under study presents clear absorption features in the visible and near-infrared range. For instance, calcium carbonate shows absorption bands in the SWIR region, the more evident near 2300 and 2350 nm due to C–O bonds of the CO_3^- radical (Bendor et al., 1999). Also, two absorption features at approx. 1400 and 1900 nm are caused by bending and stretching in the O–H bonds of free water and may be used to derive soil moisture content. Two examples are given in figure 5 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.

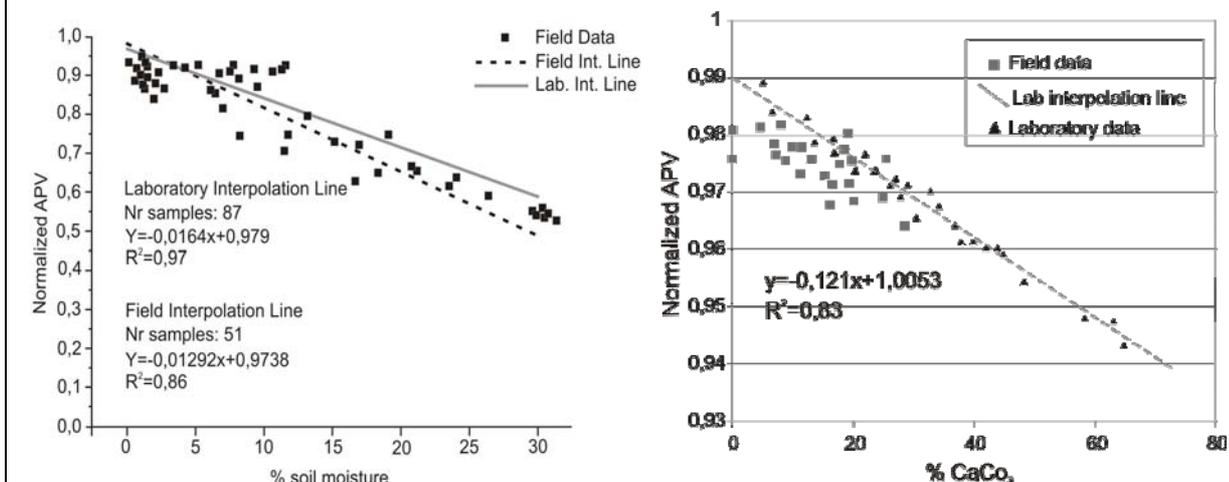


Figure 5. (a) Normalized absorption peak values (at 1460 nm) vs soil moisture (%). (b) Normalized absorption peak values (at 2436 nm) vs CaCO₃ content (%). In the laboratory, weight percentages of 5, 10, 15, 20 and 25 % of water were added to dried samples. Soil samples were also mixed with increasing amounts of a standard calcite powder (purity higher than 98%). The total concentration of CaCO₃ was calculated by adding the natural CaCO₃ content in the samples, determined by calcimetry.

Operational conditions and limits

Hyperspectral remote measurements should be taken under perfect blue sky to ensure good illumination conditions and hence a high signal-to-noise ratio. Also, since the SIM.GA sensor is on board of an ultra-light aircraft FOLDER, wind speed should be low to reduce the movement of the aircraft and avoid excessive image distortion. Flight and field campaigns should be organized in March-April or September-October during the harvest/ploughing period to collect spectral data over bare cropland soils. Optimal surface conditions occur when croplands are in seedbed conditions (low roughness), with a very low amount of vegetation residues and low water content. Compared to other geophysical techniques, hyperspectral remote sensing can only measure the first few millimetres of the surface and therefore can not monitor vertical gradients in soil properties.

References

Reference

- 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.
- Ben-Dor, E., J.R. Irons and J.F. Epema. 1999. Soil Reflectance. In: *Remote sensing for the Earth Sciences* (vol. 3): Manual of Remote Sensing (3rd ed). Renee, RA and RA Ryerson (Ed), Wiley, p. 111-188.
- 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.
- 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 Environmental Quality*, 34: 1422-1434.
- 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.
- 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
- 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.
- 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.
- 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
- 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., 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.
- 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.

6. Conclusions

The present deliverable concerns the last task of the DIGISOIL's WP1. During this study, we started to analyze the requirements of the proposed system according to the sensors, processing methodologies and applications.

This synthesis was performed according to the state of the art on the different geophysical techniques able to contribute to the Digisoil's objectives. Each (5) technique has been reviewed in terms of measurement principles, state-of-the-art and applications, strengths and limitations, and availability of commercial and non-commercial sensors.

Synthetic tables were compiled to roughly identify the integrated system that will be tested in the next field experiments.



**Scientific and Technical Centre
ARN Division**

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