Soil Surface Water Content Estimation by Full-Waveform GPR Signal Inversion in the Presence of Thin Layers

Julien Minet, Sébastien Lambot, Evert C. Slob, and Marnik Vanclooster

November 1, 2012

Abstract

We analyzed the effect of shallow thin layers on the estimation of soil surface water content using full-waveform inversion of off-ground ground penetrating radar (GPR) data. Strong dielectric contrasts are expected to occur under fast wetting or drying weather conditions, thereby leading to constructive and destructive interferences with respect to the surface reflection. First, synthetic GPR data were generated and subsequently inverted considering different thin-layer model configurations. The resulting inversion errors when neglecting the thin layer were quantified, and then, the possibility to reconstruct these layers was investigated. Second, laboratory experiments reproducing some of the numerical experiments configurations were conducted to assess the stability of the inverse solution with respect to actual measurement and modeling errors. Results showed that neglecting shallow thin layers may lead to significant errors on the estimation of soil surface water content (Δθ > 0.03 m³/m³), depending on the contrast. Accounting for these layers in the inversion process strongly improved the results, although some optimization issues were encountered. In the laboratory, the proposed full-waveform method permitted to reconstruct thin layers with a high resolution up to 2 cm and to retrieve the soil surface water content with an rmse less than 0.02 m³/m³, owing to the full-waveform inverse modeling. These results suggest that the proposed GPR approach is promising for field-scale mapping of soil surface water content of non-dispersive soils with low electrical conductivity and for instances when soil layering is encountered.

*Manuscript received May 4, 2009, revised August 3, 2009. This work was supported in part by the Belgian Science Policy Office in the frame of the Stereo II programme - project STR 00100 (HYDRASENS) and in part by Fonds de la Recherche Scientifique, Belgium.

†J. Minet and M. Vanclooster are with the Department of Environmental Sciences and Land Use Planning, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium. (e-mail: julien.minet@uclouvain.be; marnik.vanclooster@uclouvain.be).

‡S. Lambot is with the Department of Environmental Sciences and Land Use Planning, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium and also with the Agroscope (ICG-4), Institute of Chemistry and Dynamics of the Geosphere, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany. (e-mail: sebastien.lambot@uclouvain.be; s.lambot@i-geosphere.de)

§E. C. Slob is with the Department of Geotechnology, Delft University of Technology, 2628 RX Delft, The Netherlands (e-mail: E.C.Slob@tudelft.nl).

*Digital Object Identifier 10.1109/TPRS.2009.2031907
1 Introduction

At the field scale, evaluating the soil water content spatial variability is an important issue for many research and engineering applications [1]. For instance, in catchment hydrology, as the soil surface water content determines the partitioning of precipitation into run-off and infiltration under specific weather conditions, disregarding the spatial variability of the soil water content can lead to erroneous predictions in field run-off and, further, in discharge estimation of the whole catchment [2]. Usual soil water content measurement techniques at the field scale are invasive methods, like gravimetric sampling or time domain reflectometry (TDR). Although the TDR technology has been automated to some extent, the method remains problematic for mapping large areas due to the local measuring support of the TDR probe [3]. On the other hand, airborne and spaceborne remote sensing methods have been proven to be effective tools for estimating soil surface water content over larger areas, with either passive microwave radiometry or active radar instruments [4]. However, major limitations with current remote sensing techniques are the unknown within-pixel heterogeneity and the usually resulting poor agreement with calibrating and gravimetric sampling [5-9]. Hence, no absolute relation between the backscattered signals from synthetic aperture radar (SAR) and the soil water content exist, necessitating site-specific calibrations [10]. In particular, remote sensing radar systems are highly affected by soil roughness, due to the relatively high frequencies used in SAR systems, such that many studies have also addressed that problem [11]. Radar sensing is also affected by high apparent electrical conductivity values when not taken into account [12].

All of these studies show that there is a strong need to bridge the scale gap between airborne/spaceborne radar techniques and small scale characterization of soil water content. In particular, noninvasive techniques are required to characterize soil water content at the intermediate field scale and with a spatial resolution in the order of 1 m.

In that respect, ground penetrating radar (GPR) techniques are specifically suited for field scale characterization and imaging [4, 13, 14]. Several GPR techniques have been developed for a large variety of applications. For the particular case of identifying soil surface water content, two GPR approaches are commonly used. First, soil surface water content can be derived from the ground wave propagation velocity [15-18]. The ground wave is the signal that travels directly from source to receiving antenna through the soil surface. However, the technique presents a number of drawbacks, including the following: 1) the required contact between the antennas and the soil; 2) the identification of the ground wave; which may be ambiguous or even impossible in some conditions; 3) the presence of ambiguous guided waves when near-surface layering is present [19, 20]; and 4) the limited adequacy of the used straight-ray approximation for modeling electromagnetic wave propagation [21]. The second approach is the surface reflection coefficient method, which uses off-ground radar configurations [22-25]. The soil surface dielectric permittivity is derived from the Fresnel reflection coefficient, which is determined from the ratio between the amplitude of the reflection at the soil surface and the one obtained for a calibrating perfect electric conductor (PEC). However, this method still remains mostly unused nowadays for real-time mapping applications, mainly due to the requirement for such calibration [26]. The same concept is, however, commonly used in airborne and
spaceborne radar remote sensing [27]. A more recent advance has been developed by Lambot et al. [28] for the particular case of off-ground and zero-offset GPR. The method resorts to full-waveform forward and inverse modelings of the GPR signal, which inherently maximize information retrieval capabilities from the radar data. The model is based on a 3-D solution of Maxwell's equations for wave propagation in multilayered media and complex, frequency-dependent scalar transfer functions describing antenna effects, including antenna-soil interactions. Specific inversion strategies have been developed for the retrieval of soil surface dielectric permittivity and correlated water content [26].

However, Lambot et al. [26] showed that the presence of thin soil layers may significantly affect the results if not properly accounted for. Thin layers with strong dielectric contrasts may occur preferentially in coarse materials, but not limited to, either in wet soils subject to fast evaporation or dry soils subject to precipitation. To some extent, thin layers may also originate from pedo-genetic processes [29]. The soil surface is the most active layer in terms of water dynamics, as it is directly exposed to the varying atmospheric conditions, while the soil water dynamics in the subsurface is more stable. The effect of thin layers has been addressed in a limited number of studies, in the area of remote sensing with SAR [6, 7] and GPR [19, 20, 30, 31]. Currently, due to the single-frequency remote-sensing sensors, vertical contrasts of soil moisture can only be estimated using soil hydrodynamic modeling, for which knowledge of the soil hydraulic properties is required. Field or watershed scale hydraulic parameters are often derived from soil texture information using pedotransfer functions [32], but the soil parameterization schemes remain inadequate due to their inability to incorporate the natural heterogeneity of soils and the lack of detailed soil property maps. In contrast, ultra-wideband GPR particularly provides depth-dependent information and has, thereby, the potential to reconstruct thin layers.

In this paper, we analyzed the effect of thin layers on the retrieval of soil surface water content from zero-offset, normal incidence, and proximal GPR and we addressed the reconstruction of these layers by full-waveform inversion. First, numerical experiments were performed for the following reasons: 1) to investigate the well-posedness of the inverse problem when thin layers are accounted for; 2) to quantify the errors resulting from the homogeneous medium assumption; and 3) to compare different inversion strategies to deal with the reconstruction of these thin layers. Then, laboratory experiments were conducted in order to corroborate the statements inferred from the synthetic experiments and analyze the stability of the inverse problem with respect to real measurement and modeling errors. GPR measurements were made above a two-layered medium set up in a sandbox with 50 different model configurations. This paper is important for every applications where surface soil water content is estimated from remote or proximal radar data.

2 Materials and Methods

2.1 Theory

The GPR system that we used consists of a vector network analyzer (VNA) connected to an ultrawideband monostatic (zero-offset transmitter and receiver) horn antenna placed off the ground. For this configuration, the following equa-
tion is applied to filter out the antenna effects [28]:

$$S_{11}(\omega) = H_i(\omega) + \frac{H(\omega)G_{xx}^+(\omega)}{1 - H_f(\omega)G_{xx}^+(\omega)}$$  \hspace{1cm} (1)$$

where $S_{11}(\omega)$ is the quantity measured by the VNA, $H_i(\omega)$ is the antenna return loss, $H(\omega)$ is the antenna transmitting-receiving transfer function, $H_f(\omega)$ is the antenna feedback loss, $G_{xx}^+(\omega)$ is the transfer function of the air-subsurface system modeled as a multilayered medium, the so-called Green’s function, and $\omega$ is the angular frequency. The Green’s function represents an exact solution of the 3-D Maxwell’s equations for electromagnetic wave propagation in multilayered media. The consideration of a 3-D model is essential to take into account spherical divergence (geometric spreading) in GPR wave propagation. Solutions of Maxwell’s equations for wave propagation in 3-D multilayered media are well known [33]. We derived this specific Green’s function using a recursive scheme to compute the transverse electric and magnetic global reflection coefficients of the multilayered medium in the spectral domain [28, 34]. The transformation back to the spatial domain is performed by evaluating numerically a semi-infinite, complex integral. A specific procedure was applied for a fast and accurate evaluation of that integral [35], which inherently contains singularities.

The electromagnetic properties (i.e., the dielectric permittivity and the electrical conductivity) of the multilayered medium are retrieved by a full-waveform inversion of the Green’s function. This inversion can be done in the frequency domain, where the wave is actually modeled and measured, or in the time domain. For the time domain analysis, the generated and modeled frequency domain Green’s functions are first transformed in the time domain using the inverse Fourier transform. The inverse problem is formulated in the least-squares sense and the objective function is accordingly defined as follows:

$$\phi(b) = (g_{xx}^+ - g_{xx}^b)^T (g_{xx}^b - g_{xx}^+)$$ \hspace{1cm} (2)$$

where $g_{xx}^+$ and $g_{xx}^b$ are, respectively, the generated and the modeled Green’s function vectors (arranged versus frequency) and $b$ is the vector of parameters to be estimated, i.e., electromagnetic properties and dimensions of the multilayered medium. This objective function is minimized by means of the global multilevel coordinate search algorithm [36] combined sequentially with the classical Nelder-Mead simplex algorithm [37]. The reader is referred to the study by Lambot et al. [26, 28] for additional details on this model and the optimization procedure for signal inversion.

### 2.2 Numerical Experiments

The objective of these numerical experiments is to investigate and compare different inversion strategies for the estimation of thin layer properties and examine the well-posedness of the inverse problem. Synthetic radar datasets were generated assuming a two-layered model. Inversions were performed in both the frequency and time domains, assuming either the correct two-layered configuration or a simplified one-layered model. This resulted in four different inversion strategies.

The GPR frequency bandwidth used for these synthetic experiments ranges from 0.8 to 2.6 GHz, with a frequency step of 6 MHz. The antenna phase center
\( S \) was situated at 0.5 m in air above the soil surface (see Fig. 1). Different thicknesses \( h_1 \) were considered for the top thin layer \((0.005, 0.01, 0.02, 0.04\) and \(0.08\) m, as used in [26]), while the bottom layer was an infinite half-space. We chose 13 different volumetric water contents for each layer, evenly ranging from 0 to 0.36. This resulted in 845 two-layered model configurations \((13 \times 13 \times 5)\).

![2 layered configuration](image)

**Figure 1**: Model configuration used for the numerical experiments and inverse modeling flowchart.

The relations between the soil water content and its electromagnetic properties were described, respectively, by (1) the model of Ledieu et al. [38] to derive soil dielectric permittivity from water content:

\[
\theta = a\sqrt{\epsilon} + b \tag{3}
\]

with \( a = 0.1264 \) and \( b = -0.1933 \) for a specific sandy soil, and by (2) the model of Rhoades et al. [39] to relate soil electrical conductivity to water content:

\[
\sigma = (c\theta^2 + d\theta)\sigma_w + \sigma_s \tag{4}
\]

where the parameters were set to \( c = 1.85, d = 3.85 \times 10^{-2}, \sigma_w = 0.075 \text{ Sm}^{-1} \) and \( \sigma_s = 5.89 \times 10^{-3} \text{ Sm}^{-1} \). These parameters were determined in the laboratory for that specific sand subject to different water contents and salinities. Both dielectric permittivity and electrical conductivity are thus related to the soil water content by these specific relationships.

As the generated data are created with a two-layered model, the inversion procedure taking into account this layering with the two-layered model is naturally exact in contrast to the model with a one-layered configuration. However, the two-layered inversion may suffer from uniqueness and optimization problems. In addition, the inverse problem assuming a single layer is more simple, but modeling errors are then introduced. A one-layered model counts 3 parameters to invert \((\text{the electromagnetic parameters } \epsilon \text{ and } \sigma \text{ and the antenna height } h_0)\) while 5 parameters are inverted with the two-layered model \((\epsilon_1, \sigma_1, h_1, \epsilon_2, \sigma_2)\), as the antenna height above the soil is assumed to be exactly known. The error in the estimation of a parameter is defined as the absolute value of the
difference between the true parameter and the value obtained by inversion. The estimation errors are computed for each inverted parameter.

2.3 Laboratory Experiments

The objective of the laboratory experiments was to analyze in realistic conditions the same configurations of layering as for the numerical experiments. Indeed, inverse modeling of actual data are expected to show the same divergences as a function of the different model configurations and inversion strategies, plus issues related to measurement and modeling errors. Modeling errors can be caused by the antenna calibration model, the assumption of a plane layered earth and by the fact that the models of Ledieu et al. [38] and Rhoades et al. [39] may not be the correct models to relate water content to the electric properties of the layers. A schematic representation of the laboratory experimental setup is depicted in Fig. 2. A square sandbox made of wood (with 1.50 m on each side and 0.50 m in height) was filled with two layers of sand subject to specific water contents. A PEC, namely, an aluminum sheet, was placed on the bottom of the sandbox in order to control the bottom boundary condition in the electromagnetic model. As a result, deeper laboratory materials (concrete with rebar) did not influence the measured backscattered GPR signal.

The water content in the sand layers was controlled by mixing calculated volumes of dry sand and demineralized water. Volumetric samples were then collected to determine the actual water content of the sand layers. The top thin layer (layer 1 in Fig. 2) was subject to ten different water contents, ranging from 0 to 0.270. Its thickness was set to 0.005, 0.01, 0.02, 0.04 and 0.08 m, respectively. The second layer (layer 2 in Fig. 2) was subject to a constant water content of about 0.064. Its thickness was kept constant and was equal to 0.32 m. The two sand layers were separated by a thin plastic sheet. The resulting total number of model configurations was then 50 (10 × 5). The radar antenna was at a fixed position above the sandbox, at about 30 cm above the sand surface depending on the thin-layer thickness.

![Figure 2: Laboratory experimental setup showing the GPR horn antenna above the two-layered sandbox.](image-url)
We used an ultrawideband stepped frequency continuous wave radar system using a Vector Network Analyzer (VNA) (ZVRE, Rohde & Schwarz, Münich, Germany) combined with an off-ground monostatic horn antenna. The antenna system was a linear polarized double-ridged broadband horn (BBHA 9120 A, Schwarzbeck Mess-Elektronik, Schönaun, Germany). Antenna dimensions are 22 cm in length and 14 × 24 cm² in aperture area, and the nominal frequency range is from 0.8 to 5 GHz. The antenna was connected to the reflection port of the VNA with a high-quality N type 50-Ω coaxial cable of 2.5 m length (Sungoflex 104PEA, Huber+Suhner AG, Herisau, Switzerland). We calibrated the VNA at the antenna feed point using a 50-Ω OSM (Open, Short, Match) calibration kit (ZVZ21-N, Rohde & Schwarz). The frequency-dependent complex ratio $S_{11}$ between the returned and the emitted signal was measured sequentially at 301 stepped frequencies from 0.8 to 2.6 GHz, with a frequency step of 6 MHz. Measured signals from the laboratory ($S_{11}$) were first transformed in Green’s function using (1) and previously determined transfer functions [40]. These measured Green’s functions were then used for inverse modeling in order to retrieve the parameters of the two-layered medium.

Sampling for volumetric water content determination using the oven drying method at 105°C for at least 48h was performed using different sampling volumes for the different thin-layer thicknesses. Laboratory-made metallic rings were used to collect the sand samples, with heights corresponding to the thin-layer thicknesses. The bulk density of the sand was found to be 1.39 g/cm³. Five samples were taken for each mix of sand and water, accounting in total for 50 water content measurements.

The same models as the ones used in the numerical experiments (1- and 2-layered models) were used for performing the inversions in both the frequency and time domains, except that the width of the second layer ($h_2$) was inverted in the laboratory experiments as the second layer is not a half-space medium anymore. Furthermore, two other inversion scenarios were performed: 1) a two-layered model in the time domain ignoring the reflections from the PEC situated at the bottom of the sandbox (TIME 2L⁺), and 2) a one-layered model in the time domain with inversions focused on the surface reflection (TIME L-M). In the TIME 2L⁺ inversion scenario, the final reflection from the PEC is avoided by focusing on a time window defined between 0 and 5 ns, and hence, it is naturally not accounted for in the layered model. This simplified model permits a decrease in the number of unknowns in the inversion from seven to six ($\varepsilon_1$, $\sigma_1$, $h_1$, $\varepsilon_2$, $\sigma_2$ and $h_0$). In the TIME L-M inversion scenario, the GPR signal is reduced to the sand surface reflection in the time domain. In this case, the soil model reduces to a half-space and the number of parameters to estimate is two, i.e., the surface dielectric permittivity $\varepsilon_1$ and the antenna height $h_0$. This highly simplifies the inverse problem and inversions can be performed using local optimization, in our case with the Levenberg-Marquardt algorithm. This last inversion strategy has shown to be practical and suitable in different field conditions [41].
Figure 3: Error on the dielectric permittivity of the first layer $\Delta \varepsilon_1$ for each top layer thickness $h_1$, with inversions performed with the two-layered model in the frequency domain. Results from the numerical experiments.
3 Results and Discussion

3.1 Numerical Experiments

3.1.1 Inversion With the Correct Two-Layered Model

Fig. 3 shows the observed errors on inverted $\varepsilon_1$ with the correct two-layered model for the 845 considered model configurations, presented as a function of the water content of the two layers and the thickness of the thin layer. Although the model used for inverse modeling is the same as the one used for generating the synthetic radar data (i.e., a two-layered model), some discrepancies between the true and optimized parameters can be observed. For cases where there is no contrast between the two soil layers (i.e., homogeneous medium), the error is almost always null and, in all cases, negligible. For very thin layers (Fig. 3(a) and (b)), the amplitude of the error generally increases with the dielectric contrast between the two layers. For thicker layers, the error is more randomly distributed as a function of the contrast. In general, errors decrease with increasing top layer thickness. The statistical behavior of the distribution of this error follows a strong negative exponential (i.e., the mean error is 1.0059 while the median error is only 0.0107) with the majority of the errors that are null.

The observed behaviors in the error can be partly attributed to the maximum resolution that can be theoretically achieved with the considered frequency range, namely, 0.8-2.6 GHz. The range resolution is usually assumed as one quarter of the average wavelength. In dielectric media, the wavelength is decreasing proportionally to the square root of the dielectric permittivity. Hence, the range resolution increases with water content. In our case, for an average relative dielectric permittivity of nine and for a central frequency of 1.8 GHz, the range resolution is found to be 1.4 cm. This corroborates the results in Fig. 3, as significant errors are mostly observed for thinner layers. It is worth noting that in a series of cases, the observed errors for sub-resolution characterization may be also negligible. For cases where the layer thickness is larger than the range resolution, still some errors can be observed. These errors are to be attributed to optimization issues, as the global minimum of the objective function (equal to zero) could not be reached by the optimization algorithm, and the fit was therefore not perfect.

For instance, Fig. 4 represents the synthetic and fitted Green’s functions, in both the frequency and time domains, for the largest error observed for $\varepsilon_1$ (for $h_1 = 0.01$ m, $\theta_1 = 0.36$, and $\theta_2 = 0$). Although the fit is relatively good, the error in the estimated parameter is quite large. It was not possible to find a better solution with the optimization algorithm, even subject to a relatively large number of iterations (> 10000). In the time domain, the reflections from the two layer interfaces cannot be clearly distinguished, as a result of the sub-resolution conditions. However, a slight misfit can still be observed beyond the main reflection peak. The oscillations in the time domain are artifacts of the inverse Fourier transform when information is available only in a relatively limited frequency range.

Figs. 5 and 6 show 2-D slices of the 5-D objective function for a number of parameter pairs considering, respectively, a top thin layer of 1 cm (under range resolution) and 8 cm (above range resolution). In the case of the thinner
Figure 4: Generated and inverted Green’s functions for a first layer water content of 0.36, a second layer water content of 0 and a first layer thickness of 0.01 m, depicted in frequency (a) and time domain (b). Results from the numerical experiments.

layer (Fig. 5), the global minimum for the thin layer dielectric permittivity ($\varepsilon_1$) is well defined in all parameter planes. However, some correlation and weaker sensitivity can be observed in the $\varepsilon_1 - h_1$ plane (Fig. 5(c)), which is disadvantageous for its estimation. The objective function topography with respect to the electrical conductivity for both layers is always flat, which denotes a poor sensitivity of the model to these parameters. It is also worth noting that the objective function topography is relatively complex, including oscillations and local minima. In the 5-D objective function, this is therefore expected to strongly affect the optimization procedure. In Fig. 6, we observe that, compared to the 1-cm layer case, the global minimum of the objective function is better defined with increased sensitivity of the model to all parameters (except for $\sigma_1$ and $\sigma_2$). This was expected as the radar data contain sufficient information to ensure a unique estimate of the layer parameters.

3.1.2 One-Layer Model Inversions and Effect of the Contrast

Fig. 7 shows the error in the estimated dielectric permittivity of the top layer ($\Delta \varepsilon_1$) with respect to the dielectric contrast between the two layers, expressed in terms of water content contrast. Results are presented for both inversion strategies in the frequency domain, assuming the correct two-layered model and the simplified one-layered model, respectively. Errors for null contrast are not presented in this graph as they are negligible.

For the two-layered model [Fig. 7(a)], there is no clear relationship between the contrast and the error on the dielectric permittivity. We can observe the presence of numerous outliers with high error values that are out of the whisker ranges. These outliers are expected to originate from nonconvergence of the optimization algorithm.
Figure 5: Response surfaces of the objective function logarithm $\log_{10}(\phi)$ in the (a) $\varepsilon_1$-$\varepsilon_2$, (b) $\varepsilon_1$-$\log_{10}(\sigma_1)$, (c) $\varepsilon_1$-$h_1$, (d) $\varepsilon_1$-$\log_{10}(\sigma_2)$, (e) $\varepsilon_2$-$h_1$ and (f) $h_1$-$\log_{10}(\sigma_1)$ parameter planes. These objective functions are plotted for the particular case depicted in Fig. 4, i.e., $h_1 = 0.01$ m, $\theta_1 = 0.36$, and $\theta_2 = 0$. True parameter values are represented by the white star marker.
Figure 6: Response surfaces of the objective function logarithm log_{10}(\phi) in the (a) \( \varepsilon_1 \cdot \varepsilon_2 \), (b) \( \varepsilon_1 \cdot \log_{10}(\sigma_1) \), (c) \( \varepsilon_1 \cdot h_1 \), (d) \( \varepsilon_1 \cdot \log_{10}(\sigma_2) \), (e) \( \varepsilon_2 \cdot h_1 \) and (f) \( h_1 \cdot \log_{10}(\sigma_1) \) parameter planes. These objective functions are plotted for the case where \( h_1 = 0.08 \) m, \( \theta_1 = 0.36 \), and \( \theta_2 = 0 \). True parameter values are represented by the white star marker.
Figure 7: Error distribution of the dielectric permittivity of the top layer, as a function of the contrast in water content between the layers. The box extent shows interquartile range (i.e., the range between the first and the third quartiles), while the median of the error distribution is represented by the horizontal line that cuts the box. Whiskers length is 1.5 times the vertical length of the boxes. Outlier error values are displayed as crosses outside of the whiskers. Comparison between inversions performed with the two-layered (a) and the one-layered (b) models in the frequency domain. Results from the numerical experiments.
Table 1: Number of convergent configurations for time and frequency domains and for the one-layered and the two-layered inversion models in the Numerical Experiments.

<table>
<thead>
<tr>
<th></th>
<th>Error on dielectric permittivity</th>
<th>Error on the first layer thickness</th>
<th>( \Delta \varepsilon_1 ) and ( \Delta h_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \Delta \varepsilon_1 \leq 0.1 )</td>
<td>( \Delta \varepsilon_1 \leq 0.5 )</td>
<td>( \Delta h_1 \leq 0.005 ) m</td>
</tr>
<tr>
<td>Frequency domain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-layer</td>
<td>104 (12.3%)</td>
<td>268 (31.7%)</td>
<td></td>
</tr>
<tr>
<td>2-layer</td>
<td>485 (57.4%)</td>
<td>624 (73.8%)</td>
<td></td>
</tr>
<tr>
<td>Time domain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-layer</td>
<td>150 (17.7%)</td>
<td>324 (38.3%)</td>
<td></td>
</tr>
<tr>
<td>2-layer</td>
<td>342 (40.5%)</td>
<td>426 (50.4%)</td>
<td></td>
</tr>
</tbody>
</table>

For the one-layered model [Fig. 7(b)], a positive relationship between the median values of the errors on \( \varepsilon_1 \) and the contrast can be observed. In that respect, it is worth noting that the errors are shown on a logarithmic scale and, therefore, the observed trend is significant, in particular for lower contrasts. As expected, the errors for the one-layered model are larger compared to the exact model inversion (modeling errors are introduced). Large contrasts between the layers lead to significant errors as a result of constructive and destructive interferences in subresolution conditions [26] [e.g., the two reflection peaks in the time domain cannot be distinguished in Fig. 4(b)]. In case of inversion with the exact model configuration, these interferences are inherently properly accounted for. When the top layer is thick enough compared to the wavelength, the reflections at the interfaces are well separated in time and, as shown by Lambot et al. [26], the one-layered model performs well for estimating the surface dielectric permittivity.

### 3.1.3 Comparison Between Model Inversion Routines

Table 1 shows the number of convergent configurations for both time and frequency domains and for the one-layered and the two-layered inversion models. Values number the convergent cases among the 845 model configurations and percentages refer to the proportions with respect to the total. A case is counted as convergent according to the error on the dielectric permittivity of the first layer and on the first layer thickness, as classified in Table 1. This last parameter can obviously only be determined with the inversions performed with the two-layered model.

The largest convergence rate is encountered with the inversion performed in the frequency domain for the two-layered model. The one-layered model clearly shows less convergent cases than the two-layered model. This is obvious since the one-layered model does not take into account the presence of electromagnetic contrast between the layers. All the same, inversions with the one-layered model (three parameters) are about 6 times faster in terms of computation time compared to the two-layered model inversions (five parameters). Comparing time- and frequency-domain analysis, for the one-layered model, better
results are obtained in the time domain than in the frequency domain, while it is not the case with the two-layered model. As time- and frequency-domain data are equivalent in terms of information content, these differences are to be attributed to the different model sensitivities with respect to the surface dielectric permittivity. Hence, this results in different shapes for the objective function topography, with more or less well-defined global minima. Differences may also partly originate from the relation between a specific objective function topography and the used optimization algorithm (GMCS-NMS), for which the same parametrization among the scenarios was kept.

3.2 Laboratory Experiments

3.2.1 Comparison With Ground-Truth Measurements

Fig. 8 shows the soil surface relative dielectric permittivity estimated from GPR data inversion as a function of the ground-truth soil volumetric water content for six different inversion scenarios. The symbols discriminate the values according to the five top layer thicknesses that were set in the sandbox for each water content. The dotted vertical line at \( \theta_v = 0.064 \) indicates the soil water content of the bottom layer that was determined by inverse modeling of the measured GPR signal above the bottom layer without the top thin layer. In that case, the inverse problem is expected to result in accurate estimates \[28\]. For each scenario, the relationship between \( \varepsilon_{GPR} \) and \( \theta_v \) is fitted assuming the model of Ledieu et al. \[38\] (Eq. 3).

The relationships between the first-layer dielectric permittivities and water content are quite well defined and consistent for all scenarios. In general, the fits are relatively good, except for higher water contents. It is worth noting that the measured water contents are expected to be also subject to significant uncertainties, due to the following: 1) the difficulty to properly sample such thin layers with a loose material and 2) the sampling scale (5-10 cm\(^2\)), which is different from the GPR characterization scale (order of 2500 cm\(^2\)), and this is in relation to the inherent heterogeneity. The mean confidence interval at 95\%, calculated for each mix of humid sand, is equal to 0.010. Outlier permittivity values can be observed and are either very low [FREQ 2L, Fig. 8(a), TIME 2L, Fig. 8(c) and TIME 2L*, Fig. 8(e)] or very high [TIME L-M, Fig. 8(f)] compared to expected values from the fitted model of Ledieu et al. \[38\]. For overestimated values, their origin may be theoretically twofold \[26\]. First, they can result from a too high electrical conductivity, as this parameter is neglected in the TIME L-M inversion. Second, they can originate from constructive interferences in presence of thin layers in the near surface. In our case, as we observe that these discrepancies occur only for thin layers (especially 2 cm), they are believed to come from constructive interferences. In addition, the electrical conductivity is expected to be low for a sand wetted with demineralized water.

It is worth noting that the values obtained with the largest layer thickness \( (h_1 = 8 \text{ cm}) \) are very similar among the inversion scenarios. Indeed, with such layer thickness, the reflection from the second layer interface does not affect the first reflection within the given range resolution. These results for the thicker layer are therefore expected to be the most accurate.

Table 2 shows parameters \( a \) and \( b \) of the model of Ledieu et al. \[38\], the root mean square error (rmse) between the fitted and measured dielectric per-
Figure 8: Soil surface relative dielectric permittivity estimated from GPR data inversion ($\varepsilon_{1,GPR}$) as a function of soil volumetric water content $\theta_v$. Dielectric permittivities are depicted with different symbols according to the top layer thickness. The model of Leduc et al. [38] is fitted on the observed data (solid line). The dotted vertical line at value $\theta_v = 0.064$ indicates the soil water content of the second layer (constant for all configurations). (a) Two-layered inversion in the frequency domain. (b) One-layered inversion in the frequency domain. (c) Two-layered inversion in the time domain. (d) One-layered inversion in the time domain. (e) Two-layered inversion in the time domain assuming no PEC as lower-halfspace. (f) Inversion in the time domain by focusing on the surface reflection only [26].
mittivities, and the coefficient of determination $r^2$ between $\theta_v$ and $\sqrt{\epsilon_{GPR}}$ for the corresponding six scenarios. Except for the TIME L-M scenario, the one-layered models show the lowest rmse and largest coefficients of determination $r^2$. In general, better results are obtained with the one-layered model as the dimensionality of the inverse problem is lower. Similarly, the TIME 2L* scenario also presents relatively low rmse as, in that case, only six parameters are to be inverted for (see also Fig. 8(e)). For scenario TIME L-M, a large rmse and a low $r^2$ are obtained, arising from the presence of numerous outliers. These outliers come from low model adequacy, as it does not account for the constructive and destructive interferences produced by the two interfaces.

The two-layered models show intermediate performances in terms of rmse and $r^2$, as a trade-off between model adequacy and inverse problem complexity. A smart parameterization of the inverse algorithm is required because of the high non-linearity of the model to its parameters. In that way, reducing the number of parameters to optimize considerably makes the inverse problem easier. For instance, the best performance of the model TIME 2L* compared to FREQ 2L and TIME 2L comes simultaneously from the limited number of parameters (passing from seven to six) and high model adequacy.

Fig. 9 shows the GPR-derived top layer thickness ($h_2$) as a function of the real thickness measured using a millimetric ruler. Inversions were performed with the TIME 2L* scenario, which generally produces the best results. The observed errors are relatively small, ranging from a few millimeters up to two centimeters. The most accurate estimations are obtained for the 4-cm-thick layer. For thicknesses lower than 2 cm, discrimination between the two interfaces cannot always be achieved, because, in that case, the range resolution limit is reached. As it was shown in the numerical experiments, this range resolution is about 1.5 cm for a center frequency of 1.7 GHz and a dielectric permittivity of nine. For larger thicknesses, the larger errors are to be attributed to the negative correlation between the layer thickness and its dielectric permittivity (see Fig. 6). In presence of measurement and modeling errors, the minimum region of the objective function becomes even flatter, which results in larger uncertainties in the estimation of these parameters.
3.2.2 Effect of the Contrast

As it was already observed in the numerical experiments, high contrasts between the two soil layers lead, in general, to larger errors in the estimation of the surface dielectric permittivity with one-layered models. In Fig. 8, the dotted vertical lines correspond to the water content of the bottom layer, which is 0.064, and therefore, at this line, the contrast between the two layers is zero. Accordingly, we clearly see that, for this water content, the dispersion of the dielectric permittivities around the model of Ledieu et al. [38] for the one-layered models is minimal. Indeed, the two-layered configuration then reduces to a one-layered configuration. We can also observe that, for the zero contrast, inversion results are better for the one-layered inversion, because of the lower dimensionality of the inverse problem. When the contrast between the two layers increases, the inadequacy between the one-layered model and the reality (two layers) increases, which results in proportionally larger errors.

For the two-layered inversions, in general, error also increases with the contrast. As in that case the forward model is correct, these discrepancies are to be attributed either to optimization issues or a badly defined global minimum of the objective function. As this was not observed for the numerical experiments, we believe that the problem mainly originates from the objective function topography, which inherently becomes flatter when measurement errors are present.

Fig. 10 shows the absolute value of the difference between the measured and the fitted top layer water content as a function of the contrast between the layers for the TIME L-M inversion scenario. This figure emphasizes the positive correlation between the contrast and the error on the GPR-derived soil properties, as already observed in Fig. 8. It is important to mention here that this inversion strategy (i.e., a time domain inversion focusing on the surface
reflection only) appears to be very robust in general when dealing with field data [41, 42]. In that respect, these laboratory experiments show that one-layered inversions even perform relatively well in the presence of thin layers with low contrasts. For high contrasts, however, accounting for thin layers is necessary.

3.2.3 Comparison Between One- and Two-Layered Models for Estimating the Surface Dielectric Permittivity

Fig. 11 shows the absolute difference in dielectric permittivities of the top layer between the one- and two-layered inversions, as a function of the ground-truth top-layer water content. This figure allows the evaluation of the resulting error when a one-layered model is assumed for a two-layered medium. Gray areas in the graph delineate two dielectric permittivity error thresholds, namely, $\Delta \varepsilon_1 \leq 1$ and $\Delta \varepsilon_1 \leq 3$. Dielectric permittivities are depicted with different symbols according to the first-layer thickness.

All the values that are above the two thresholds pertain to configurations with small layer thicknesses, i.e., $h_1 = 0.5$ cm, $h_1 = 1$ cm and, to a lesser extent, $h_1 = 2$ cm. On the other side, configurations whose first-layer thickness are $h_1 = 8$ cm systematically lead to low differences (i.e., very close dielectric permittivity estimations are obtained with the two model configurations). As expected, this shows that a top layer thick enough compared to the wavelength can be characterized from the surface reflection only. It also appears that errors slightly increase with higher water contents, denoting the dielectric contrast effect.
3.2.4 Effect of the frequency bandwidth sampling

Information content in the radar data increases with increasing bandwidth. An important asset of using VNA technology is that the bandwidth is fully controllable and, in particular, ultrawideband GPR can be set up. The only limitation is the operating frequency range of the antenna. However, in some cases, other factors may also affect the utilizable bandwidth. In our case, in the laboratory conditions, lower frequencies were subject to ambiguous reflections from extraneous objects present in the laboratory, given the limited size of the sandbox. Measurements for higher frequencies were also influenced to some extent by the inherent heterogeneities present in the sandbox, including slight surface roughness. For instance, Fig. 12 shows that the GPR measured signal becomes of poorer quality for the lowest and highest frequencies of the 0.8-2.6 GHz bandwidth we used, and in particular below 1 GHz and above 2 GHz. The correct two-layered model used for the inversion could reproduce remarkably well the measured signal in both the frequency and time domains.

In order to determine the optimal frequency bandwidth for the top layer dielectric permittivity estimation, we performed inversions in different limited frequency ranges selected from the full range, and we assumed the correct two-layered model (FREQ 2L). Fig. 13 presents boxplots of the errors in the top-layer water content estimation for inversions performed in eight different frequency ranges. The errors are defined as the bias between the GPR soil water content derived from the dielectric permittivities using the model of Ledieu et al. [38] with a given parametrization and the soil water content determined by volumetric sampling. Start frequencies were set to either 0.8 or 1 GHz and stop
frequencies to 2, 2.2, 2.4 or 2.6 GHz, respectively.

Clearly, the full frequency bandwidth leads to the best estimations for the sand water content. Narrower bands lead, in particular, to outliers with significant errors. Reducing the frequency bandwidth is clearly contraindicated for estimating the soil parameters in the presence of thin layers. For layers that are thick compared to the wavelength, a bandwidth as small as 0.4 GHz still leads to accurate results [43]. Although the radar data for frequencies between 0.8 GHz and 1 GHz are contaminated by some noise, inversions without these frequencies do not lead to significantly different results. A large bandwidth is important for thin-layer retrieval because, in the frequency domain, thin layers lead to oscillations of the amplitude of the Green’s function with a large period. This period is inversely proportional to the electric thickness of the layer. It is worth noting that the problem is similar in the time domain where the limitation of the bandwidth results in less well-defined reflection (the bandwidth defines range resolution). For instance, with remote SAR acquisition, it is well known that using various frequencies enhances the extraction of information for the retrieval of soil surface moisture [10]. Moreover, D'Urso and Minacapilli [10] have shown that relatively low frequencies (L-band, 1.6 GHz) lead to better results, mainly due to the lower sensitivity to soil surface roughness and vegetation compared to higher frequencies.

4 Conclusions and Perspectives

We performed numerical experiments to evaluate the effect of shallow thin layers on the retrieval of soil electromagnetic properties by full-waveform inversion of GPR data. First, inversions of synthetic data, assuming the correct model configuration (two-layered), show some discrepancies between true and inverted dielectric permittivities. This was attributed to the complexity of the inverse
Figure 13: Error boxplots between the GPR-derived and the sampled volumetric soil water content of the top layer, for inversions with different frequency bandwidths. The box extent shows interquartile range (i.e., the range between the first and the third quartiles), while the median of the error distribution is represented by the horizontal line that cuts the box. Whiskers length is 1.5 times the vertical length of the boxes. Outlier error values are displayed as crosses outside of the whiskers.
problem (five parameters to be retrieved) and the not well-defined global minimum of the objective function for subresolution layer thicknesses. We also showed that the errors increase, as expected, with the use of a simplified one-layered model (i.e., when the true two-layered medium is not taken into account). However, provided limited contrast between the two layers, such approach remains robust owing to the low dimensionality of the inverse problem (three parameters). For relatively large contrasts (e.g., > 0.10 in terms of water content), significant errors arise with the simplified model (> 0.10 in terms of water content).

For the laboratory experiments, a good agreement was obtained between volumetric water content determined by sampling and GPR-derived dielectric permittivity. The same limitations as for the numerical experiments were observed. In addition, part of the observed discrepancies was attributed to the inherent variability of the water content within the sand layers with respect to the different measurement support, making the water content measurements not fully reliable in comparing the larger-scale GPR measurements. Retrieval of the thin-layer thickness led to relatively small errors, ranging from a few millimeters up to a maximum of two centimeters (corresponding to the range resolution).

The benefit of these experiments and analysis is that it shows both the theoretical and practical limits in terms of thin layers reconstruction using zero-offset, off-ground GPR. These results also apply to other radar remote sensing, but considering a much larger bandwidth. Depending on shallow layering conditions, different inversion strategies should be adopted (i.e., based on one- and two-layered model configurations). The proposed methods appear to be promising for soil surface water content mapping at the field scale or any applications where the nondestructive surface properties of media are to be determined.

References


