Northumbria Research Link

Northumbria University
NEWCASTLE

UniversityLibrary

Fractal-Constrained Crosshole/Borehole-to-Surface Full Waveform Inversion for Hydrogeological Applications Using Ground-Penetrating Radar

Iraklis Giannakis, Antonios Giannopoulos, Craig Warren and Anastasia Sofroniou

Abstract—**Full waveform inversion (FWI) is considered one of the most promising interpretation tools for hydro-geological applications using ground-penetrating radar. However, FWI has had limited practical uptake for several reasons: large computational requirements; an inability to reconstruct loss mechanisms of soil; and the need for a good initial starting model. We aim to address these issues via a novel FWI subject to a fractally-correlated distribution of water. Initially, the dispersive properties of the soil are expressed as a function of the water fraction using a semi-empirical model. This approach means the permittivity, conductivity, and relaxation mechanisms are all correlated, and therefore sensitivity problems between the permittivity and loss mechanisms no longer affect the performance of FWI. Subsequently, the distribution of the water fraction is constrained to follow a fractal geometry. Fractal-correlated noise is then compressed using principal component analysis (PCA) in order to further reduce the number of the system's unknowns and accelerate FWI. PCA reduces the volume and dimensions of the optimization space, and thus, initialisation is no longer necessary. Lastly, a novel measurement configuration is suggested that uses superposition with all the individual measurements in order to reduce the number of forward models that need to be executed for every iteration of FWI. These enhancements substantially reduce the computational requirements of FWI and therefore eliminate the need for high-performance computers and time-consuming algorithms. The proposed scheme has been successfully tested with several numerical case-studies, which indicate the potential of this approach to become a commercially-appealing interpretation tool for hydro-geology.**

Index Terms—**Ground Penetrating Radar (GPR), Full-Waveform Inversion (FWI), Fractals, Hydrogeology, Finite-Differences Time-Domain (FDTD), Principal Component Analysis (PCA).**

I. INTRODUCTION

GROUND-penetrating radar (GPR) is a mature geophysical technique with a diverse set of applications, ranging from glaciology [1] and planetary sciences [2], to archaeology [3] and landmine detection [4]. GPR has been established as a mainstream interpretation tool in hydrogeology [5] and has been successfully applied for permeability estimation [6],

I. Giannakis is with the School of Geosciences, University of Aberdeen, Meston Building, Kings College, Aberdeen, AB24 3FX, UK. E-mail: iraklis.giannakis@abdn.ac.uk

A. Giannopoulos is with the School of Engineering, The University of Edinburgh, Edinburgh, EH9 3FG, UK. E-mail: a.giannopoulos@ed.ac.uk

C. Warren is with the Department of Mechanical and Construction Engineering, Northumbria University, Newcastle, NE1 8ST, UK. E-mail: craig.warren@northumbria.ac.uk

A. Sofroniou is with the School of Computing and Engineering, University of West London, London, W5 5RF, UK. E-mail: anastasia.sofroniou@uwl.ac.uk

mapping the water table [7], delineating clay layers [8], characterizing aquifers [9], deriving the root zone soil moisture [10] and investigating groundwater-dependent ecosystems [11].

In order to address the diversity of the hydrogeological applications that GPR is used for, numerous surveying strategies and corresponding processing pipelines have been suggested over the years [12]. One of the most popular is the crosshole/borehole-to-surface measurement configuration, aimed at interpolating hydrogeological properties between boreholes [13], [14], [15]. The interpretation of these data is generally based on the geometrical ray theory [16], which utilises the first arrivals and the maximum amplitudes [17], [18] (subject to regularizations and constraints [19], [20], [21]) in order to reconstruct the dielectric properties of the investigated medium. Although ray-based tomography is practical and computationally efficient, it suffers from significant short-falls primarily associated with an inherent lack of resolution [16], [22], [23].

Full-waveform inversion (FWI) addresses these issues by utilising all the available information within the received signal. This allows the overall resolution of the tomography to be greatly increased [24], and the ability of GPR to characterise loss mechanisms is enhanced [16]. FWI was initially suggested for seismic prospecting [25] and has been extensively investigated since then [26]. The FWI of GPR has a wide range of applications (using both crosshole [27] and on-ground measurement configurations [28]) which vary from estimating the radius of cylindrical objects [29] and mapping the chloride content of concrete [30], to characterizing aquifers [13], [14].

Recent theoretical developments have further established FWI as a reliable interpretation tool for GPR acquisitions and especially for crosshole/borehole-to-surface hydrogeological surveys [15]. In particular, source deconvolution [31], [32] and using the ratio of two electromagnetic field parameters [33] have been successfully applied for removing the effects of the unknown source wavelet. Moreover, in an effort to reduce the computational requirements, 2.5D forward solvers and 2D to 3D transformations have been proposed in order to replace costly 3D simulations [34]. Furthermore, the presence of numerous local minimal in the optimization space has been tackled via ray-based initial models [16], using global optimizers [35], [36], and by gradually increasing the selected bandwidth as the optimization progresses [37]. Lastly, a further increase in resolution was made possible by addressing

sharp boundaries and promoting "blocky" solutions using total variation regularization constraints [38].

Despite these advancements, there are still several issues and pitfalls that need to be addressed before FWI can become commercially appealing to GPR practitioners and hydrogeologists. The most important of which is the very demanding –and often unattainable– computational requirements. Numerous evaluations of the forward model combined with time-consuming Jacobian derivations and memory-intense matrix operations require high-performance computers to operate for prolonged periods of time. Another major drawback of FWI is that it fails to provide an accurate estimation of the conductivity [15] and –to our knowledge– no attempt has been made to characterize the full dispersive properties of the investigated medium. Moreover, the final results are sensitive to the starting model [15] and therefore FWI is not applicable in the absence of sufficient initialisation. Lastly, FWI greatly relies on the inversion parameters (regularizations etc.) [15], and a parameter selection based on trial-and-error is not attainable due to the high computational requirements for each FWI execution.

This paper presents research which aims at addressing the aforementioned issues through a novel FWI, constrained to a compressed representation of the fractal distribution of the water fraction of the subsurface. The suggested scheme utilises a superpositioned forward model combined with semi-empirical formulas [39], [40] for deriving the extended Debye properties of the soil. This results in an efficient FWI scheme capable of characterising the full dielectric properties of the soil without the need for regularization and initialisation. Finally, due to the imposed fractal-constrains, our scheme is orders of magnitude faster compared to traditional FWI, which makes a step change towards the commercialisation of FWI in hydrogeology and borehole geophysics.

II. METHODOLOGY

The proposed FWI tries to address inherent shortfalls in traditional FWI via the following novel attributes: a superpositioned forward solver that reduces the number of forward models needed to be executed in each iteration; a semi-empirical dielectric model for soils [39], [40] that overcomes the sensitivity discrepancies between electrical permittivity and conductivity; and a fractal-constrained inversion that reduces the dimensions of the optimisation space and therefore accelerates FWI without the need for initialization and memory-intensive Jacobian derivations.

A. Super-positioned forward model

The forward solver of the proposed FWI is a second order finite-difference time-domain (FDTD) method [41]. FDTD is particularly attractive for GPR applications due to its time-domain nature and its ability to implement dispersive media [43]. Extended Debye relaxation mechanisms are accurately incorporated into the model using the polarization density method [44] while the boundaries of the grid are effectively truncated via the time-synchronised convolutional perfectly matched layer boundary conditions [45].
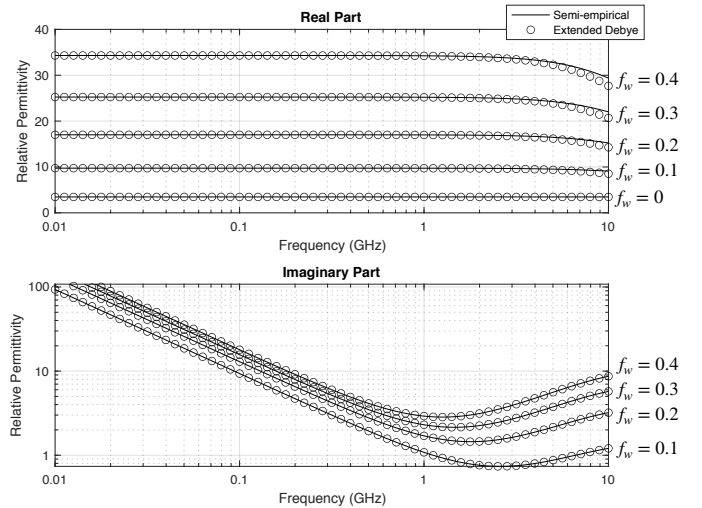

Fig. 1. Real and imaginary parts of the electrical permittivity (shown by solid lines) calculated using the semi-empirical model described in equations (1) - (9) for different volumetric water fractions f_w [39], [40]. The extended Debye approximations using the analytical approach described in [48] is plotted with circles. Notice that for $f_w = 0$, the imaginary part becomes zero and therefore is omitted from the plot.

Crosshole/borehole-to-surface configurations consist of multiple transmitters and receivers placed inside two parallel boreholes and along the surface between them. For each transmitter, the resulting electromagnetic fields are simultaneously recorded by the receivers. The full dataset (i.e. all the recorded signals for every transmitter) is then used in order to find the optimum model -using computational electrodynamics- that gives rise to a similar response. FDTD –and most of the numerical methods in computational electrodynamics– is capable of calculating the received signal in many locations with just one execution [41], with additional receivers adding very minimal extra computational burden to forward modelling. On the other hand, multiple transmitters require multiple executions, which consequently increase the overall computational requirements of FWI. In order to overcome this, we use the superposition principle which states that the net response caused by multiple transmitters equals with the sum of the responses for each transmitter individually [42]. Therefore, the received signal using different transmitters can be summed and the resulting dataset will be equal to the dataset that would have been measured if all the transmitters were triggered simultaneously. Consequently, the forward model needs to be executed only once, using simultaneously-triggered impressed sources at the positions of the transmitters. Via this approach, the computational requirements of FWI are greatly reduced since forward modelling is typically one of the most time-consuming and memory-intense components of FWI.

B. Semi-empirical dielectric model for soils

FWI can not reliably reconstruct the distribution of conductivity since the latter is particularly sensitive to the initial model as well as the inversion parameters [15]. To tackle this, the proposed FWI utilises a semi-empirical model, initially developed by [39] and subsequently revised by [40], to express
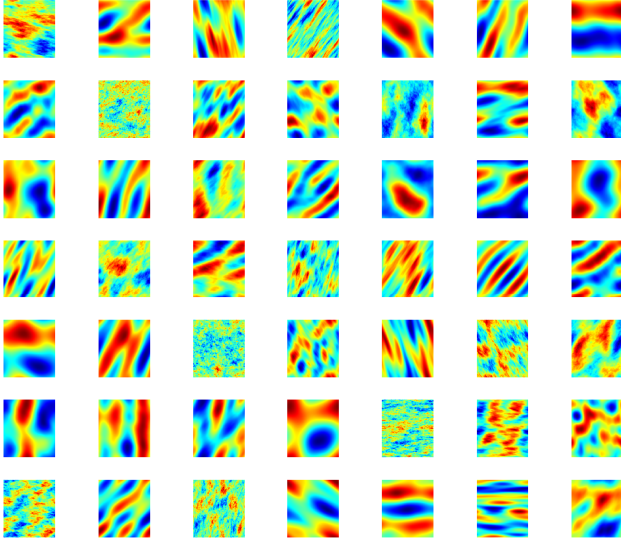
Fig. 2. A random sample of the various features that can be generated using anisotropic fractal correlated noise (15). The scale of the axis is omitted since fractals are scale invariant [61].
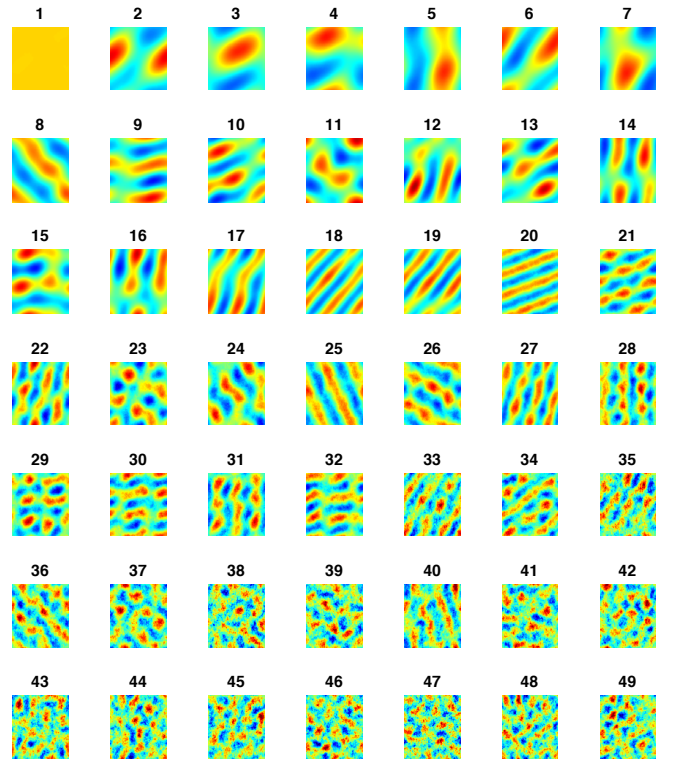

Fig. 3. The first 49 (out of r) principal components of a training set consisting of 2000 randomly generated anisotropic 2D fractals. The title of each subfigure corresponds to the ith number of the principal component q_i. Notice that the dominant (first) principal components relate to the low-wavenumber elements of the model.

the full dispersive properties of the soil with respect to the water fraction. Consequently, the parameter that the proposed FWI tries to recover is the water fraction which corresponds to a specific extended Debye model. The rationale behind this approach is that –for hydrogeological applications– the permittivity, conductivity and dipolar relaxation mechanisms are highly correlated with each other and the most dominant factor affecting their behaviour is the presence or absence of water [39], [40], [46], [47].

Equations (1) - (9) describe the semi-empirical model that relates the dispersive dielectric properties to the soil properties [39], [40]

$$\epsilon'_{1.4-18 \ GHz} = \left(1 + \frac{\rho_b}{\rho_s}(\epsilon_s^\alpha - 1) + f_w^{\beta'}\epsilon_w'^\alpha - f_w\right)^{1/\alpha} \quad (1)$$

$$\epsilon'_{0.3-1.3GHz} = 1.15\epsilon'_{1.4-18GHz} - 0.68 \quad (2)$$

$$\epsilon'' = f_w^{\frac{\beta''}{\alpha}}\left(\epsilon_w'' + \frac{\sigma_f}{\omega\epsilon_0}\frac{(\rho_s - \rho_b)}{\rho_s f_w}\right) \quad (3)$$

$$\epsilon_s = (1.01 + 0.44\rho_s)^2 - 0.062 \quad (4)$$

$$\beta' = 1.2748 - 0.519 \cdot S - 0.152 \cdot C \quad (5)$$

$$\beta'' = 1.33797 - 0.603 \cdot S - 0.166 \cdot C \quad (6)$$

$$\epsilon_w = \epsilon_{w,\infty} + \frac{\epsilon_{w,s} - \epsilon_{w,\infty}}{1 + j\omega t_{0,w}} \quad (7)$$

$$\sigma_{1.5-18GHz} = -1.645 + 1.939\rho_b - 2.25622 \cdot S + 1.594 \cdot C \quad (8)$$

$$\sigma_{0.3-1.3GHz} = 0.0467 + 0.2204\rho_b - 0.411 \cdot S + 0.6614 \cdot C \quad (9)$$

where f_w is the water volumetric fraction, $\epsilon'_{1.4-18GHz}$ and $\epsilon'_{0.3-1.3GHz}$ are the real parts of the permittivity for the frequency range of 1.4 - 18 GHz and 0.3 - 1.3 GHz respectively, and ϵ_0 is the electric permittivity of free space. The imaginary part of the permittivity is defined by ϵ'', ρ_s is the sand particles density (g/cm^3), ρ_b is the bulk density

of the soil (g/cm^3), ϵ_s is the relative permittivity of the sand particles, and $\alpha = 0.65$ is an experimentally derived constant. The parameters S and C are the sand and clay mass fractions respectively ($0 \le S, C \le 1$, $S + C = 1$), and $\epsilon_w = \epsilon_w' + j\epsilon_w''$ is the complex electric permittivity of water. The relaxation time of the water is defined by $t_{0,w}$ (s), $\epsilon_{w,s}$ is the relative permittivity of the water for zero frequency, and $\epsilon_{w,\infty}$ is the relative permittivity of the water for infinite frequency. Lastly, the constants $\sigma_{f1.4-18GHz}$ and $\sigma_{f0.3-1.3GHz}$ are linearly proportional to the conductivity σ for the frequency range of 1.4-18 GHz and 0.3-1.3 GHz respectively. The soil described in (1)-(9) is considered non-magnetic and therefore the relative magnetic permeability is one.

The semi-empirical model described in (1) - (9) cannot be directly implemented in FDTD, and thus an extended Debye function is used to approximate it for the frequency range of interest [43], [48]. The analytical method described in [48] is used for the approximation. No iterative optimization is required for this methodology and therefore it does not add to the computational burden of FWI. The resulting extended Debye function ϵ_d is given by [48]

$$\epsilon_d = \epsilon_{d,\infty} + \frac{\sigma_d}{j\omega\epsilon_0} + \frac{\epsilon_{d,s} - \epsilon_{d,\infty}}{1 + j\omega t_{0,d}} \quad (10)$$

$$\sigma_d = f_w^{\frac{\beta''}{\alpha}}\sigma_f\frac{\rho_s - \rho_b}{\rho_s f_w} \quad (11)$$
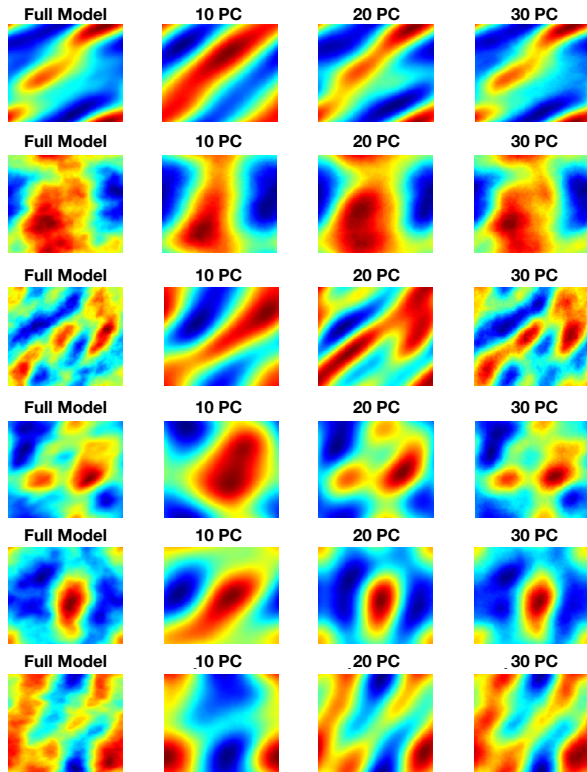
$$t_{0,d} = t_{0,w} \quad (12)$$

Fig. 4. Six random fractal models (not part of the training set **T**) are approximated using a linear combination of 10, 20, and 30 principal components (PC).

$$\epsilon_{d,s} = \epsilon'(0) \tag{13}$$

$$\epsilon_{d,\infty} = \epsilon'(0) - f_w^{\frac{\beta''}{\alpha}} \left(\epsilon_{w,s} - \epsilon_{w,\infty} \right) \tag{14}$$

The most important parameter in the semi-empirical model -especially for frequencies below 1 GHz- is the water volumetric fraction [48]. The dielectric properties of water are related to temperature T [49], [50] and can be chosen based on borehole measurements. For the current paper, we have chosen $T = 20^o C$ which corresponds to $t_{0,w} = 9.23$ ps, $\epsilon_{w,s} = 80.1$, and $\epsilon_{w,\infty} = 4.9$. Regarding the soil properties, typical soils have $\rho_s = 2.66 \ g/cm^3$ and $\rho_b = 2 \ g/cm^3$ and small variations around these values have small effects to the overall complex permittivity [48]. Moreover, although C and S play an important role for frequencies greater than 1.5 GHz (by changing the ratio between free and bounded water [48]), they have minor effects on the frequencies most often used in hydrogeological applications [48], and thus are set to the average $S = C = 0.5$. For frequencies below 1 GHz, the clay content affects the dielectric properties indirectly due to the high porosity of clays (i.e. high f_w) and not due to the clay particles themselves. Even Cole-Cole phenomena, predominantly present in clay media, are masked by the overall effects of the volumetric water fraction for the frequency range of interest [48].

Figure 1 illustrates the complex electrical permittivity for $T = 20^o C$, $\rho_s = 2.66 \ g/cm^3$, $\rho_b = 2 \ g/cm^3$, $S = C = 0.5$,
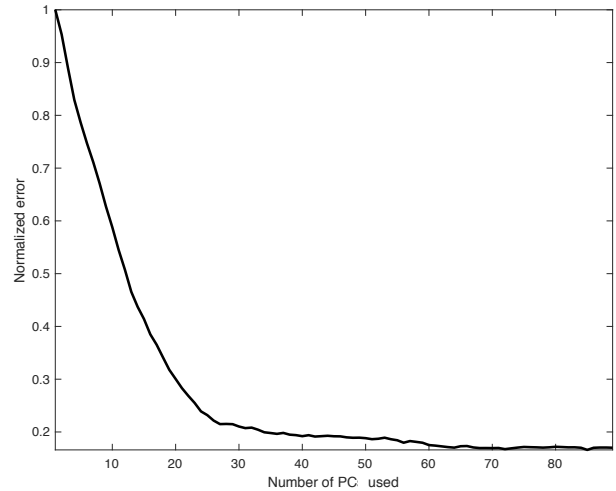

Fig. 5. The mean absolute difference of 200 randomly generated fractals and their corresponding approximation using different number of principal components (PC).

and for $f_w = 0 - 0.4$ using the semi-empirical model and corresponding extended Debye approximations. Notice that via this approach, both the real and the imaginary part (conductivity and dipolar losses) of the complex electrical permittivity can be inferred and subsequently incorporated in FDTD using the volumetric water fraction f_w as the only input.

C. Fractal-constrained FWI

Typical FWI is an ill-posed problem that suffers from instabilities, non-uniqueness, and local-minimal [38]. To that extent, good initialisation and regularisation methods, such as Tikhonov and total variation constraints, are necessary for providing a stable and reliable solution [38]. In addition, resolution-requirements result in a large number of model parameters that need to be fine-tuned. This leads to time-consuming and memory-intensive Jacobian and Hessian calculations that increase the overall computational requirements of FWI. In order to tackle the aforementioned issues, we suggest a novel FWI scheme in which the water fraction of the soil is forced to follow a compressed anisotropic fractal distribution with an arbitrary fractal dimension. This choice is based on a large volume of evidence in the literature that suggest that hydrogeological, soil, and environmental parameters can be sufficiently described using fractal correlated noise [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61]. In particular, realistic synthetic reservoir models were generated using fractals in [51]. Landscapes and various environmental data were also accurately modelled based on fractal distributions in [52], [61]. Lastly, fractals have also been extensively used for modeling the physical properties of soils [53], [55], [57], [58], [59], [60] such as hydraulic conductivity and porosity [54], [56]. Based on the above, it is evident that under normal geological conditions, soil properties are expected to follow a correlated distribution that can be sufficiently described using fractals. Nonetheless, this does not hold true for non-stochastic, deterministic distributions, often encountered in
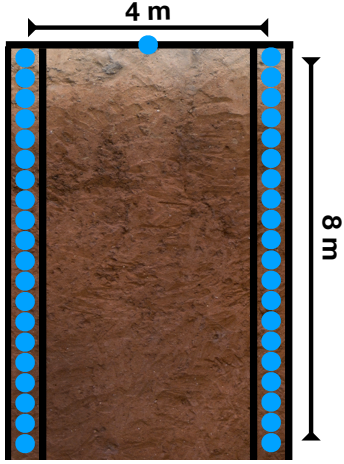
Fig. 6. Measurement configuration used for the numerical case studies examined in this paper. Twenty Tx/Rx are placed in each borehole and one Tx/Rx is placed on the surface between the boreholes. All antennas act both as transmitters and receivers and they are triggered simultaneously in the FDTD simulation. Subsequently the direct wave is removed by time-gating.

anthropogenic environments. In those cases, the performance of the fractal-constrained FWI will be limited since FWI will try to approximate the deterministic model with a compressed orthonormal basis derived based on fractal correlated noise.

Additionally, as shown later in this section, PCA was proven to be very effective in compressing anisotropic fractals. This results in a significant reduction (by orders of magnitude) of the model parameters that need to be optimized. The reduced model parameters, combined with the semi-empirical model used for the soil properties, greatly accelerate FWI without the need for regularization methods and initialization.

Anisotropic fractal correlated noise for N dimensions can be expressed as [43], [48]

$$F(x_1...x_n) = \text{rot}_\theta \left(\mathfrak{F}^{-1} \left(R(k_1...k_n) \cdot \left(\sum_{i=1}^{N} a_i k_i^2 \right)^H \right) \right) \tag{15}$$

where F is the fractal correlated noise, $\mathfrak{F}^{-1}$ is the sign for the inverse Fourier transform, R is an Nth dimensional Gaussian noise, x_i is the ith dimension of the fractal, k_i is its corresponding wavenumber, a_i is a weight parameter that generates anisotropic fractals [51], rot is a rotational operator that rotates the image at an angle θ, and H is a linearly proportional parameter to the fractal dimension of F [61]. Figure 2 illustrates a sample of the features that can be generated using (15). Changing H, θ, and a_i results in an infinite number of different scenarios with a plethora of patterns ranging from layer structures and tilted inhomogeneities, to distinct structures of various shapes and sizes (see Figure 2).

An arbitrarily-generated 2D-fractal distribution of water fraction is defined as $\mathbf{Q}(f_w) \in \mathbb{R}^{x \times z}$, where x and z are the dimensions of the model. Consequently, the total number of parameters that need to be optimized by FWI is $r = x \times z$. The dimensions of the optimization space can be significantly reduced by compressing $\mathbf{Q}(f_w)$ using PCA [62], [63]. Initially, the 2D matrix $\mathbf{Q}(f_w)$ is reshaped into the vector $\mathbf{q}(f_w) \in \mathbb{R}^r$.

We then create a coherent training set of 2000 randomly generated 2D fractals $\mathbf{T} = [\mathbf{q}_1, \mathbf{q}_2...\mathbf{q}_{2000}] \in \mathbb{R}^{r \times 2000}$. The fractal parameters H, θ and a_i are randomly selected subject to a uniform distribution. Via this approach, a coherent, equally-distributed and representative training set is generated, capable of adequately resolving the investigated feature space. The principal components of the training data $\mathbf{T}$ are stored in a square matrix $\mathbf{P} \in \mathbb{R}^{r \times r}$ consisting of orthogonal vectorized images with decreasing co-variance. Figure 3 illustrates the first 49 principal components of the training set $\mathbf{T}$. The first principal components correspond to low-wavenumber components (background) while the later principal components describe finer structures. PCA suggests that every vectorized fractal $\mathbf{q}(f_w)$ of the training data, can be expressed as a linear combination of all principal components i.e. $\mathbf{q}(f_w) = \mathbf{Pw}$, where $\mathbf{w} \in \mathbb{R}^r$ is a column vector with the same dimensions as $\mathbf{q}(f_w)$. Choosing a subset of the principal components $\mathbf{P} \in \mathbb{R}^{r \times u}$ ($u < r$) results in an over-determined system $\mathbf{q}(f_w) \approx \mathbf{Pv}$, where $\mathbf{v}(f_w) \in \mathbb{R}^u$ is the vector that contains the compressed coefficients of $\mathbf{q}(f_w)$. The latter can be calculated using the well-known least squares formula $\mathbf{v}(f_w) = \left(\mathbf{P}^T \mathbf{P} \right)^{-1} \mathbf{P}^T \mathbf{q}(f_w)$. Consequently, the compressed fractal distribution of the water fraction can be calculated via $\mathbf{L}(f_w) = \mathbf{Pv}(f_w)$, where $\mathbf{L}(f_w) \in \mathbb{R}^{x \times z}$. Anisotropic fractals have been proven to be very compressible and using just the first 30 principal components is proven adequate for sufficiently reconstructing any randomly generated anisotropic fractal (see Figure 4 and 5). Essentially we are deriving an orthonormal basis $\mathbf{P}$ fine-tuned for fractal geometries. Meaning that a small subset of $\mathbf{P}$ is adequate to reconstruct fractal-correlated noise. For non-fractal geometries, more principal components are needed. For a fully accurate reconstruction, the whole matrix $\mathbf{P}$ should be used which will result to a $\mathbf{v}$ vector with length that equals to $\mathbf{w}$ i.e. the size of the model (zero compression).

The vector $\mathbf{v}(f_w)$ contains the 30 parameters of the model that need to be optimised in the next step. These 30 parameters represent a compressed anisotropic fractal distribution of the water fraction, which corresponds to a fractal distribution of an extended Debye model that describes the full dielectric behaviour of the soil based on the semi-empirical model previously presented. Therefore, the investigated medium can be fully described –subject to the aforementioned constraints– using just 30 scalars with known boundaries (based on the training set). Here we need to highlight that fractal-constrain is a soft constrain similar to smoothness and total variation constrains; and does not force the model parameters to be a specific shape or type. It just makes the solutions biased to scale-invariant and self-similar geometries i.e. follow a fractal distribution (see Fig. 2).

D. Optimization method

The proposed FWI tries to find the best vector $\mathbf{v}(f_w)$ that minimizes the summation of the absolute difference between the actual and the simulated measurements. The transmitters are all triggered simultaneously, and therefore the cost function
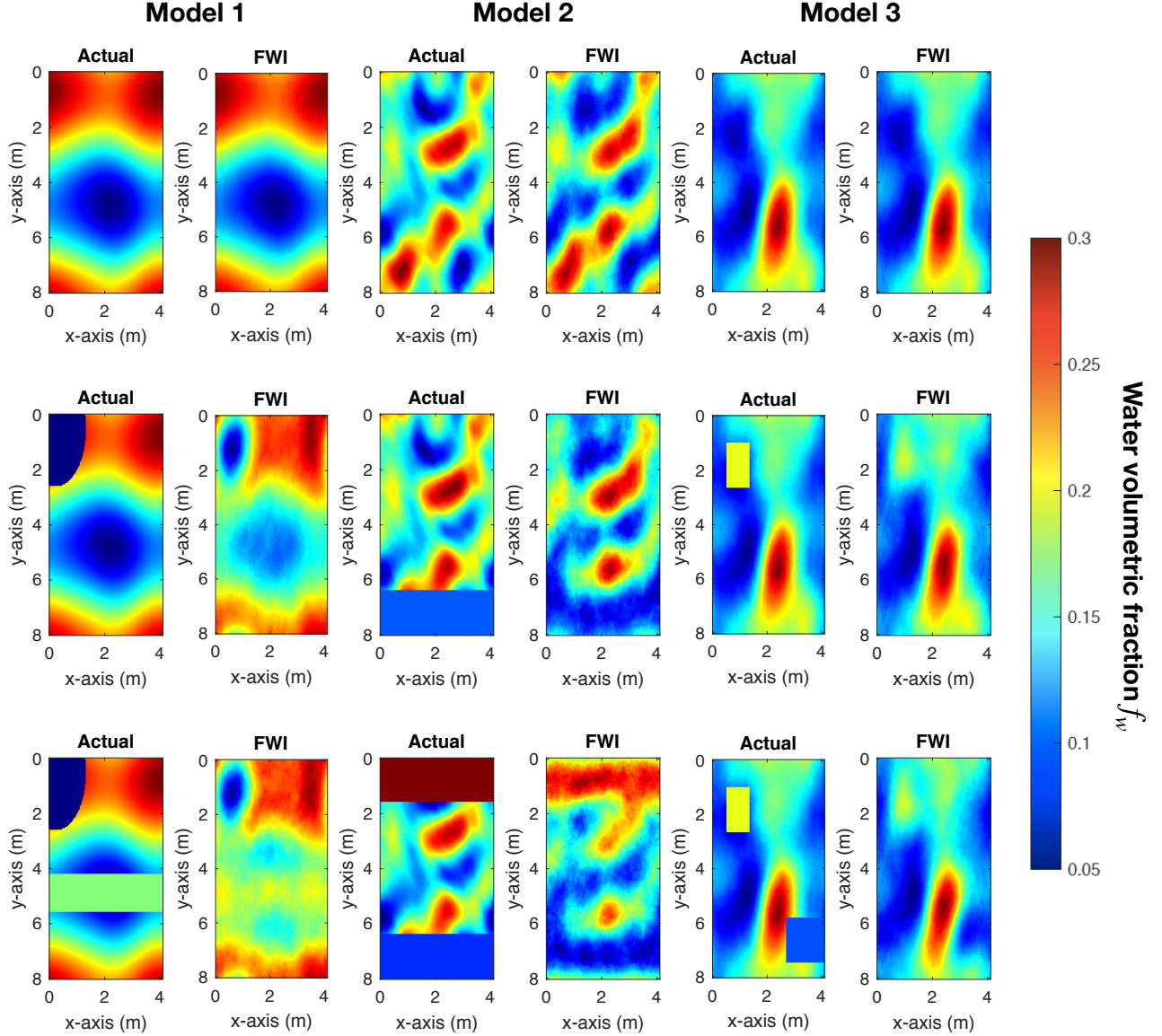
Fig. 7. Three numerical case studies are used to demonstrate the capabilities of the proposed interpretation scheme. The first row corresponds to the investigated fractal models. In the second and third row, various deterministic non-fractal targets are included in the model.

$S(\mathbf{v})$ needs a single FDTD execution in order to be evaluated

$$S(\mathbf{v}) = \sum_{Tx} \sum_{Rx} \sum_{t} |E^o_{Tx,Rx,t} - E^s_{Tx,Rx,t}(\mathbf{v}) * A(t,\mathbf{v})| \quad (16)$$

where Tx is the transmitter index, Rx is the receiver index, t is the observation time, E^o are the observed measurements, and E^s are the simulated ones. The simulated traces are generated using an arbitrary source-wavelet that deviates from the actual pulse leading to errors and instabilities [31]. The effects of the unknown source can be formulated via a convolutional term $A(t,\mathbf{v})$ that acts upon all measurements [64]

$$\sum_{Tx} E^{rs}_{Tx,Rx,t}(\mathbf{v}) = \left(\sum_{Tx} E^s_{Tx,Rx,t}(\mathbf{v}) \right) * A(t,\mathbf{v}) \quad (17)$$

where E^{rs} is the received trace subject to the correct source-wavelet. The convolutional term $A(t,\mathbf{v})$ is calculated by tak-

ing the average value of the deconvolution between all the observed and simulated signals

$$A(t,\mathbf{v}) = \mathfrak{F}^{-1} \left(\frac{1}{N} \sum_{Tx} \sum_{Rx} \frac{E^o_{Tx,Rx,\omega}}{E^s_{Tx,Rx,\omega}(\mathbf{v}) + e} \right) \quad (18)$$

where ω is the angular frequency, N is the total number of receivers, and e is a very small damping constant to address divisions by zero. In (18) we try to find the best wavelet-source that fits the observed data for a given vector $\mathbf{v}(f_w)$. Consequently, the function $A(t,\mathbf{v})$ is re-calculated subject to the updated $\mathbf{v}(f_w)$ as the optimisation progresses.

The optimization space of $\arg\min_{\mathbf{v}} S(\mathbf{v})$ is bounded by the expected values of $\mathbf{v}(f_w)$ derived from the training set $\mathbf{T}$. Therefore, each component $v_i(f_w)$ of the compressed water fraction lies within a predefined domain $v_i(f_w) \in [l_i, u_i]$. The first step of the optimisation is a sequential independent search

for every principal coefficient in $\mathbf{v}(f_w) = [v_1, v_2....v_{30}]$. In particular, the cost function $S(v_1)$ is evaluated for five equidistant values between $[l_1, u_1]$. A spline interpolation is applied to fit the resulting points, and the optimum v_1 that minimises $S(v_1)$ is chosen. Subsequently, the cost function $S([v_1, v_2])$ is evaluated using the value of v_1 from the previous step and five equidistant values for v_2 between $[l_2, u_2]$. Similar to v_1, a spline interpolation is applied to the resulting points, and the v_2 that minimises $S([v_1, v_2])$ is chosen. This procedure continues until all the principal coefficients are set. Via this approach, the initial model $\mathbf{v}_0(f_w) = [v_1, v_2...v_{30}]$ of the optimisation is evaluated. The first step of the optimisation follows the notion described in [37], [65] where low frequency components are initially inverted in order to reconstruct the background model and avoid local minimal. In the proposed FWI, similar to [37], [65], the model is progressively reconstructed starting with the first principal coefficients (background, see Figure 3) and progressively moving to finer structures. The compressed model $\mathbf{v}_0(f_w)$ is subsequently used as an initial solution for an interior-point nonlinear optimisation algorithm as presented in [66], [67]. The interior-point is a fully bounded method that deals with non-convexity and can be applied in a straightforward manner using the function *fmincon* in MATLAB [68]. Via trial and error it was estimated that, for the majority of the case-studies examined, the interior-point method converges in ≈ 20 iterations.

III. NUMERICAL CASE STUDIES

The proposed FWI is tested in various 2D numerical case studies including different scenarios with different water distributions and fractal dimensions. Figure 6 shows the employed measurement configuration. Twenty Tx/Rx are placed in each borehole and one Tx is placed between them at the top surface. As it is described in section II.A, all the Tx are excited simultaneously in order to reduce the number of times that the FDTD forward model needs to be executed. The spatial step of the FDTD is $\Delta x = \Delta y = 0.1$ m and the time-step follows the Courant limit [41]. The excitation pulse used to generate the synthetic data is an ideal line source excited by a Gaussian modulated pulse with central frequency 80 MHz. Only the central frequency of the antenna is considered to be known during FWI. Instead of the actual pulse, the forward solver of FWI is excited by the first derivative of a Gaussian function with the same central frequency as the real pulse. The whole FWI procedure takes around 20-30 minutes on an Apple MacBook Air with a 1.6 GHz Intel Core i5 CPU and 8 GB 1600 MHz DDR3 RAM.

Figure 7 illustrates the actual and the reconstructed water distributions using the suggested fractal-constrained FWI. Three numerical case studies are investigated with different water distributions and fractal dimensions. Each model is further modified by adding non-fractal deterministic geometries in order to evaluate the performance of the suggested scheme to non-fractal targets within fractal environments. From Figure 7 it is evident, that using fractals as basis functions, makes it possible to reconstruct complex geometries in an accurate, and most importantly, time efficient manner.
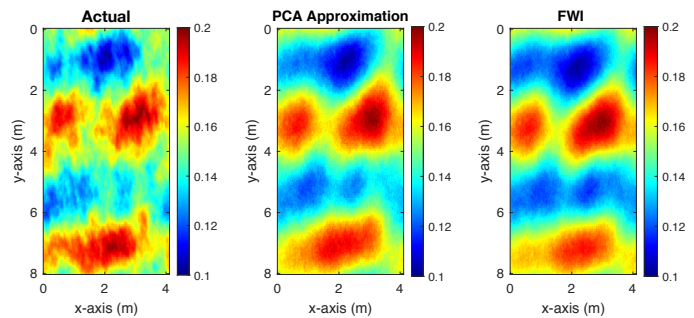

Fig. 8. First column represents the real model. Second column represents the PCA approximation using 30 PCs. The third column is the estimated water fraction using the fractal constrained FWI. Colorbar represents the water volumetric fraction f_w.

The next numerical example (Figure 8) illustrates the inherited resolution constrains due to the limited (compressed) number of PC used for the reconstruction. The performance of the proposed FWI is inversely proportional to the roughness of the investigated model. The resolution of the model is constrained primarily by the employed central frequency and the number of PC used for the approximation (see Figure 8). Increasing the resolution of FWI would require higher frequency antennas combined with additional principal components in $\mathbf{v}(f_w)$. Increasing the number of principal components –and consequently the computational requirements of FWI– using the same central frequency will have negligible effects to the overall resolution of the model. Changing the central frequency will increase the resolution but will decrease the investigation area due to the reduced penetration depth. Therefore, a change in the frequency component of the pulse needs to be combined by an equivalent scaling of the investigation area. Notice that since fractals are scale invariant, changing the central frequency and the scale of the model will have no affect on the principal components needed to approximate a given fractal.

The proposed algorithm is then tested on purely deterministic non-fractal geometries often encountered in anthropogenic environments. In situations like these, fractal-constrained FWI should be used with caution since it is inherently constrained to approximate any given geometry by its fractal equivalent. Figure 9 illustrates three case studies with simple geometrical targets. It is evident that the proposed method estimates a fractal approximation of the investigated geometry and fails to reconstruct sharp boundaries and clinical non-fractal geometries.

Lastly, we test the fractal-constrained FWI in non-stacked data corrupted with random Gaussian noise. Figure 10 illustrates the investigated case study. Data from a low dimensional fractal model were corrupted with 0%, 5% and 10% Gaussian noise. It is apparent, that Gaussian noise effects the reconstructed model by modifying the boundaries between permittivity structures; and by adding a high-wavelength noise giving the impression of a model with a higher fractal-dimension.
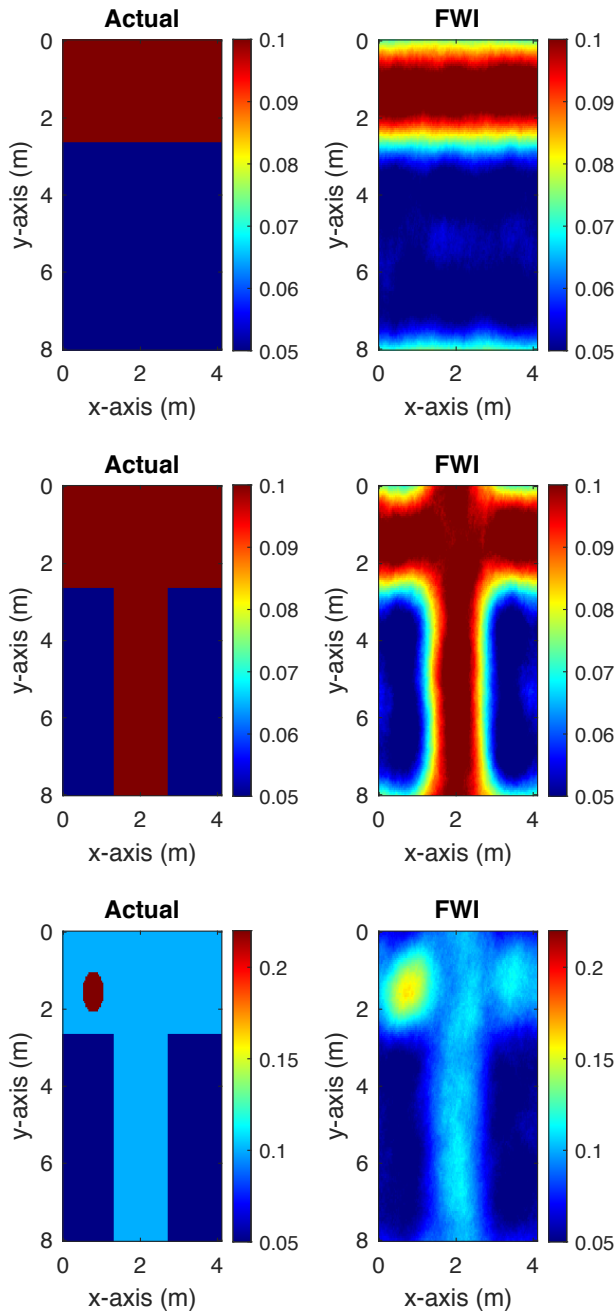
Fig. 9. Evaluation of the suggested FWI scheme on non-fractal media. The colorbar indicates the water volumetric fraction f_w.
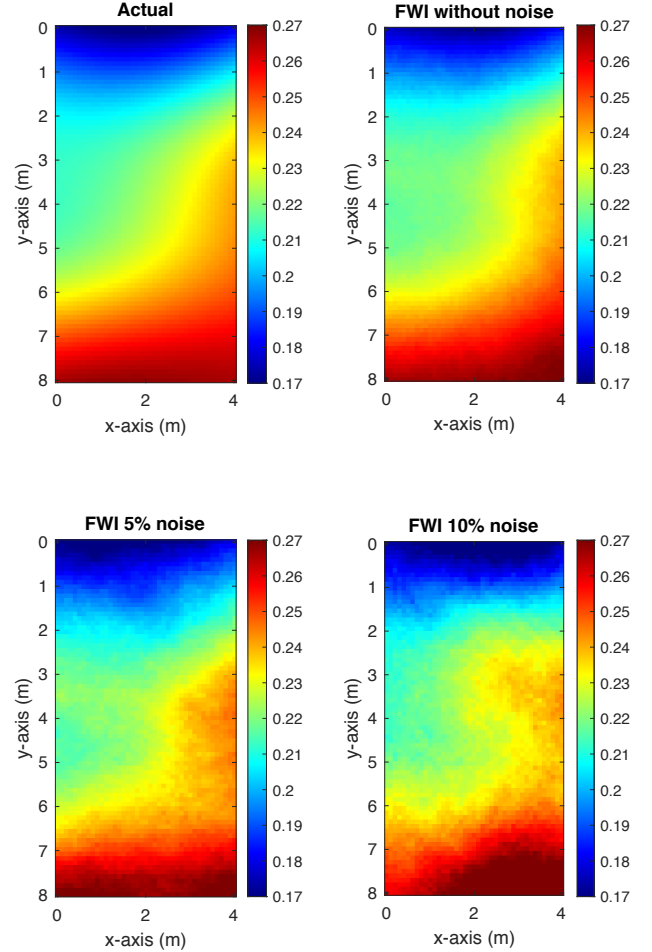

Fig. 10. Assessing the performance of the proposed scheme subject to contaminated data with 0%, 5% and 10% Gaussian noise. The colorbar indicates the water volumetric fraction.

economic forward solver, further contributes to the reduction of the overall memory and computational requirements. These novel additions make the suggested FWI commercially appealing and attainable for GPR practitioners and hydrogeologists without the need for high performance computers.

IV. CONCLUSIONS

A novel fractal-constrained full-waveform inversion (FWI) is suggested in order to reconstruct the water distribution between boreholes in a practical and efficient manner. Via a set of numerical experiments, it is demonstrated that a fractal distribution of water can be successfully reconstructed using its compressed representation based on principal component analysis. Within that context, inherent properties of fractals are utilised in order to reduce the number of the unknowns and accelerate FWI. In addition, a semi-empirical model that reduces the number of the unknowns, combined with a more

REFERENCES

[1] A. Rutishauser, H. Maurer, and A. Bauder, "Helicopter-borne ground-penetrating radar investigations on temperate alpine glaciers: A comparison of different systems and their abilities for bedrock mapping," *Geophysics*, vol. 81, pp. 1JF-Z7, 2016.

[2] C. Li, Y. Su, E. Pettinelli, [...] and H. Zhang, "The Moon's farside shallow subsurface structure unveiled by Chang'E-4 Lunar Penetrating Radar," *Science Advances*, vol. 6, no. 9, eaay6898, 2020.

[3] L. Verdonck, A. Launaro, F. Vermeulen, and M. Millett, "Ground-penetrating radar survey at Falerii Novi: a new approach to the study of Roman cities," Antiquity, vol. 94, no. 375, pp. 705–723, 2020.

[4] D. J. Daniels, "A review of landmine detection using GPR," in *Proc. of European Radar Conference*, Amsterdam, 2008, pp. 280-283, 2008.

[5] A. P. Annan, "GPR Methods for Hydrogeological Studies," In: Rubin Y., Hubbard S.S. (eds) *Hydrogeophysics*, Water Science and Technology Library, vol 50. Springer, Dordrecht, 2005.

[6] R. M., Corbeanu, G. A., McMechan, R. B. Szerbiak, K. Soegaard, "Prediction of 3D fluid permeability and mudstone distributions from ground-penetrating radar (GPR) attributes: example from the Cretaceous Ferron Sandstone member, east-central Utah," *Geophysics*, vol. 67, pp. 1495–1504, 2002.

[7] J. A., Doolittle, B. Jenkinson, D. Hopkins, M. Ulmerd, W. Tuttlee, "Hydropedological investigations with ground-penetrating radar (GPR): estimating water-table depths and local ground-water flow pattern in areas of coarse-textured soils," *Geoderma*, vol. 131, pp. 317–329, 2006.

[8] D. Gómez-Ortiz, T. Martín-Crespo, S. Martín-Velázquez, P. Martínez-Pagán, H. Higueras and M. Manzano, "Application of ground penetrating radar (GPR) to delineate clay layers in wetlands. A case study in the Soto Grande and Soto Chico watercourses, Doñana (SW Spain)," *J. Appl. Geophys.*, vol. 72, pp. 107–113, 2010.

[9] J. Doetsch, N. Linde, M. Pessognelli, A. G. Green, T. Günther, "Constraining 3-D electrical resistance tomography with GPR reflection data for improved aquifer characterization," *J. Appl. Geophys*, vol. 78, pp. 68–76, 2012.

[10] X. Liu, J. Chen, R. J., Butnor, G. Qin, X. Cui, B. Fan, H. Lin and L. Guo, "Noninvasive 2D and 3D mapping of root zone soil moisture through the detection of coarse roots with ground-penetrating radar," *Water Resources Research*, vol. 56, e2019WR026930, 2020.

[11] C. Paz, F. J. Alcalá, J. M. Carvalho and L. Ribeiro, "Current uses of ground penetrating radar in groundwater-dependent ecosystems research," *Science of the Total Environment*, vol. 595, pp. 868-885, 2017.

[12] D. J. Daniels, *Ground Penetrating Radar*, 2nd ed. London, U.K.: Institution of Engineering and Technology, 2004.

[13] A. Klotzsche, J. van der Kruk, G. A. Meles, J. Doetsch, H. Maurer and N. Linde, "Full-waveform inversion of cross-hole groundpenetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland," *Near Surface Geophysics*, vol. 8, pp. 635–649, 2010.

[14] A. Klotzsche, J. van der Kruk, G. A. Meles, and H. Vereecken, "Crosshole GPR full-waveform inversion of waveguides acting as preferential flow paths within aquifer systems," *Geophysics*, vol. 77, no. 4, pp. H57–H62, 2012.

[15] A. Klotzsche, H. Vereecken, and J. van der Kruk, "Review of crosshole ground-penetrating radar full-waveform inversion of experimental data: Recent developments, challenges, and pitfalls," *Geophysics*, vol. 84, pp. H13-H28, 2019.

[16] G. A. Meles, J. Van der Kruk, S. A. Greenhalgh, J. R. Ernst, H. Maurer, and S. A. Greenhalgh, "A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3391–3407, Sep. 2010.

[17] T. C. Johnson, P. S. Routh, W. Barrash and M. D. Knoll, "A field comparison of Fresnel zone and ray-based GPR attenuation-difference tomography for time-lapse imaging of electrically anomalous tracer or contaminant plumes," *Geophysics*, vol. 72, no. 2, pp. G21-G29, 2007.

[18] F. D. Day-Lewis, J. M. Harris, and S. M. Gorelick, "Time-lapse inversion of crosswell radar data," *Geophysics*, vol. 67, pp. 1740–1752, 2002.

[19] E. Gloaguen, D. Marcotte, M. Chouteau and H. Perroud, "Borehole radar velocity inversion using cokriging and cosimulation," *Journal of Applied Geophysics,* vol. 57, pp. 242-259, 2005.

[20] E. Gloaguen, B. Giroux, D. Marcotte and R. Dimitrakopoulos, "Pseudo-full-waveform inversion of borehole GPR data using stochastic tomography," *Geophysics*, vol. 72, pp. J43–J51, 2007.

[21] B. Giroux, E. Gloaguen and M. Chouteau, "bh_tomo—a Matlab borehole georadar 2D tomography package," *Computer Geosciences*, vol. 33, pp. 126–137, 2007.

[22] P. R. Williamson, "A guide to the limits of resolution imposed by scattering in ray tomography," *Geophysics*, vol. 56, no. 2, pp. 202-207, 1991.

[23] P. R. Williamson and M. H. Worthington, "Resolution limits in ray tomography due to wave behavior—numerical experiments," *Geophysics*, vol. 58, no. 5, pp. 727-735, May 1993.

[24] R. G. Pratt and R. M. Shipp, "Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data," *Geophysics*, vol. 64, no. 3, pp. 902–914, Jun. 1999.

[25] A. Tarantola, "Inversion of seismic reflection data in the acoustic approximation," *Geophysics*, vol. 49, pp. 1259–1266, 1984.

[26] J. Virieux and S. Operto, "An overview of full-waveform inversion in exploration geophysics," *Geophysics*, vol. 74, pp. WCC1–WCC26, 2009.

[27] S. Busch, J. van der Kruk, and H. Vereecken, "Improved characterization of fine-texture soils using on-ground GPR full-waveform inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 3947–3958, Jul. 2014.

[28] J. R., Ernst, A. G. Green, H. Maurer, K. Holliger, "Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data," *Geophysics* vol. 72, pp. J53–J64, 2007.

[29] T. Liu, A. Klotzsche, M. Pondkule, H. Vereecken, J. van der Kruk, and Y. Su, "Estimation of subsurface cylindrical object properties from GPR full-waveform inversion," *in Proc. 9th Int. Workshop Adv. Ground Penetrating Radar (IWAGPR)*, Jun. 2017, pp. 1–4

[30] A. Kalogeropoulos, J. van der Kruk, J. Hugenschmidt, J. Bikowski, and E. Brühwiler, "Full-waveform GPR inversion to assess chloride gradients in concrete," *NDTE Int.*, vol. 57, pp. 74–84, Jul. 2013.

[31] F. A. Belina, J. Irving, J. R. Ernst, and K. Holliger, "Waveform inversion of crosshole georadar data: Influence of source wavelet variability and the suitability of a single wavelet assumption," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4610–4625, Nov. 2012.

[32] F. Belina, J. Irving, J. Ernst, and K. Holliger, "Analysis of an iterative deconvolution approach for estimating the source wavelet during waveform inversion of crosshole georadar data," *J. Appl. Geophys.*, vol. 78, pp. 20–30, Mar. 2012.

[33] J-H Kim, T. Kobayashi and S. K. Lee, "Admittance inversion of GPR transmission for crosshole tomography," *Journal of Applied Geophysics*, vol. 81, pp. 57-67, 2012.

[34] A. Mozaffari, A. Klotzsche, C. Warren, G. He, A. Giannopoulos, H. Vereecken, and J. van der Kruk, "2.5D crosshole GPR full-waveform inversion with synthetic and measured data," *Geophysics* vol. 85, pp. H71-H82, 2020.

[35] O. F. Mojica and N. Kukreja, "Towards automatically building starting models for full-waveform inversion using global optimization methods: A PSO approach via DEAP + Devito," arXiv:1905.12795, 2019.

[36] F. A. Moura, S. A. Silva, J. M. de Araújo and L. S. Lucena, "Progressive matching optimisation method for FWI," *Journal of Geophysics and Engineering*, vol. 17, pp. 357–364, 2020.

[37] G. A. Meles, S. Greenhalgh, J. van der Kruk, A. G. Green, H. Maurer, "Taming the non-linearity problem in GPR full-waveform inversion for high contrast media," *Journal of Applied Geophysics*, vol. 73, pp. 174-186, 2011.

[38] D. Feng, C. Cao and X. Wang, "Multiscale Full-Waveform Dual-Parameter Inversion Based on Total Variation Regularization to On-Ground GPR Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 9450-9465, Nov. 2019,

[39] M. C. Dobson, F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes, "Microwave dielectric behavior of wet soil—Part II: Dielectric mixing models," *IEEE Trans. Geosci. Remote Sens.*, vol. GRS-23, pp. 35–46, Jan. 1985.

[40] N. R. Peplinski, F. T. Ulaby, and M. C. Dobson, "Dielectric properties of soils in the 0.3–1.3-GHz range," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 3, pp. 803–807, May 1995.

[41] A. Taflove and S. C. Hagness, *Computational Electrodynamics, the Finite-Difference Time-Domain Method*, 2nd ed. Norwood, MA, USA: Artech House, 2000.

[42] C. Balanis, *Advanced Engineering Electromagnetics*, Wiley New York, 1989.

[43] I. Giannakis, A. Giannopoulos and C. Warren, "A realistic FDTD numerical modelling framework of ground penetrating radar for landmine detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 1, pp. 37-51, 2016.

[44] I. Giannakis and A. Giannopoulos, "A novel piecewise linear recursive convolution approach for dispersive media using the finite-difference time-domain method", *IEEE Trans. Antennas Propag.*, vol. 62, no. 5, pp. 2669-2678, 2014.

[45] I. Giannakis and A. Giannopoulos, "Time-synchronised convolutional perfectly matched layer for improved absorbing performance in FDTD", *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 690-693, 2015.

[46] A. Brovelli and G. Cassiani, "Effective permittivity of porous media: a critical analysis of the complex refractive index model," *Geophysics*, vol. 56, pp. 715-727, 2008.

[47] P. Hoekstra, A. Delaney, "Dielectric properties of soils at UHF and microwave frequencies," *Journal of Geophysical Research*, vol. 79, no. 11, pp. 1699-1708, 1974.

[48] I. Giannakis, *Realistic numerical modelling of ground penetrating radar for landmine detection,* PhD Thesis Submitted at The University of Edinburgh, 2016.

[49] L. A. Klein and C. T. Swift, "An improved model for the dielectric constant of sea water at microwave frequencies," *IEEE Transactions on Antennas and Propagation*, vol. 25, no. 1, pp. 104-111, 1977.

[50] T. A. Phuong, M. R. Mahmoudzadeh Ardekani and S. Lambot, "Coupling of dielectric mixing models with full-wave ground-penetrating radar signal inversion for sandy-soil-moisture estimation," *Geophysics*, vol. 77, no. 3, pp. H33-H44, 2012.

[51] S. Al-Zainaldin, P. W. J. Glover and P. Lorinczi, "Synthetic fractal modelling of heterogeneous and anisotropic reservoirs for use in simulation studies: implications on their hydrocarbon recovery prediction," *Transport in Porous Media*, vol. 116, pp. 181-212, 2017.

[52] P. A. Burrough, "Fractal dimensions of landscapes and other environmental data," *Nature*, vol. 294, pp. 240–242, Nov. 1981.

[53] E. Perfect and B.D. Kay, "Applications of fractals in soil and tillage research: a review," *Soil and Tillage Research*, vol. 36, pp. 1-20, 1995.

[54] D. Giménez, E. Perfect, W.J. Rawls and Y. Pachepsky, "Fractal models for predicting soil hydraulic properties: a review," *Engineering Geology*, vol. 48, pp. 161-183, 1997.

[55] D. Hillel, *Environmental Soil Physics*, New York, NJ, USA: Academic Press, 1980.

[56] J. D. Jabro, W. B. Stevens, R. G. Evans and W. M. Iversen, "Spatial variability and correlation of selected soil properties in the AP horizon of a CRP grassland," *Applied Engineering in Agriculture*, vol. 26, pp. 419-428, 2010.

[57] R. B. Rezaur, H. Rahardjo and E. C. Leong, "Spatial and temporal variability of pore-water pressures in residual soil slopes in a tropical climate," *Earth Surface Processes and Landforms*, vol. 27, pp. 317-338, 2002.

[58] K. Takahashi, J. Igel, H. Preetz and M. Sato, "Influence of heterogeneous soils and clutter in the performance of ground-penetrating radar for landmine detection," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 52, no. 6, pp. 3464-3472, 2014.

[59] Q. F. Yansui Liu and M. Mikami, "Geostatistical analysis of soil moisture variability in grassland," *Journal of Arid Environments,* vol. 58, pp. 357-372, 2004.

[60] A. K. Verma, P. Kumar, W. K. Mohanty, A. Routray and B. A. Cheadle, "Modeling of reservoir property using fractal interpolation and geostatistics," *SEG Technical Program Expanded Abstracts:* pp. 2460-2464, 2014.

[61] D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics*, Cambridge, U.K.: The Press Syndicate of the Univ. Cambridge, 1992.

[62] M. C., Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1996.

[63] X. Huang, L. Wu, Y. Ye, "A review on dimensionality reduction techniques," *International Journal of Pattern Recognition and Artificial Intelligence*, Early access available, vol. 33, no. 10, 1950017, 2019.

[64] I. Giannakis, A. Giannopoulos and C. Warren, "Realistic FDTD GPR Antenna Models Optimized Using a Novel Linear/Nonlinear Full-Waveform Inversion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1768-1778, March 2019.

[65] H. Maurer, S. A. Greenhalgh and S. Latzel, "Frequency and spatial sampling strategies for acoustic crosshole full waveform inversion experiments," *Geophysics*, vol. 74 pp. WCC79-WCC89, 2009.

[66] H. R., Byrd, E. M. Hribar and J. Nocedal, "An interior point algorithm for large-scale nonlinear programming," *SIAM Journal on Optimization*, vol. 9, no. 4, pp. 877–900, 1999.

[67] H. R. Byrd, J. C. Gilbert and J. Nocedal, "A trust region method based on interior point techniques for nonlinear programming." *Mathematical Programming*, vol. 89, no. 1, pp. 149–185, 2000.

[68] https://uk.mathworks.com/help/optim/ug/fmincon.html, Jul.2020.