

Geostatistics inversion of electric field, waveform synthetic data, with synthetic a priori data of well electrical permittivity properties to GPR, using the scaled conjugate gradient

Abstract

Electromagnetic wave data are simulated in representation of waves registered by a Ground Penetrating Radar (GPR), post-stacked and migrated, by convolving the synthetic electric permittivity of a model of parallel layers of soil and a ricker wavelet. Data well information of synthetic electrical permittivity is perturbed close to the original synthetic data on electric permittivity properties associated to a model conformed with parallel layers. We define a priori probability density function on the space of models, then use Bayes theorem to combine this probability with the data misfit function into a final a posteriori probability density reflecting both data fit and model reasonableness, we incorporate the covariance matrix of the well data and synthetic data. The posteriori probability is optimized via a nonlinear scaled conjugate gradient. The input data of the scaled conjugate gradient algorithm are the well permittivity electric data, the Ground Penetrating Radar source wavelet (Ricker) and the synthetic electromagnetic wave data that represents the waves recorded by a GPR post-stacked and migrated are contaminated with 5% noise and without noise. The SCG algorithm is initialized to an average constant value of the well data.

The results obtained from the inversion are very close to the initial real synthetic values of the electric permittivity model of parallel layers of the soil, that we wanted to recover, via this proposed inversion technique.

Keywords: simulation, electromagnetics wave propagation, electric permittivity

Volume 2 Issue 3 - 2018

Julio C Morfe,¹ Ana I Bermudez,¹ Herbert Rendon²

¹GrupoLi, Medellin, Colombia

²Funvisis, Caracas, Venezuela

Correspondence: Julio C Morfe, GrupoLi, Medellín, Colombia, Email peac63@hotmail.com

Received: May 21, 2018 | **Published:** June 21, 2018

Introduction

Geophysical data inversion deals with the problem of a quantitative inference on Earth model parameters from data observation. One sound approach to this problem is what is called the Bayesian inference method, where a priori information on the Earth parameters is known in terms of its probability distribution function and new data is sought to improve the knowledge of Earth Model parameters in terms of a better constrained a posteriori probability distribution functions of those very same parameters. In this work we apply the Bayesian approach to electric permittivity data where we believe that a priori information can reasonably be derived from in situ petrophysical/field measurements. Therefore, due to observations are subject to uncertainty, these inferences are inherently probabilistic. Also, due to the epistemic error associated to uncertainties in the model parameters, this call for a framework which includes both a model for the random noise associated to the capturing of the data and the one associated to the lack of certain knowledge on the fabrics on the Earth media. Therefore, our task is to find an improved Earth model that fits new collected data, but also complies with our prejudice of what it should be a reasonable model, avoiding such “truly” unrealistic models that violate our a priori prejudices, other data, or theoretical considerations. One strategy for eliminating such unreasonable models is to define an a priori probability density on the space of models, then use Bayes theorem to combine this probability with the data misfit function into a final a posteriori probability density function that reflects both the data and a reasonable model.

Reflection of electromagnetic waveform

Formulation of the problem

In our procedure, we perform a Bayesian inverse calculation to estimate relative electric permittivity. The Bayesian solution of the problem is the a posteriori probability density function on model space, which we denote by $\sigma(\bar{Z})$. This probability density function is the convolution product of two terms, the likelihood function $L(\bar{Z})$, which defines what is meant for a model to fit the data, and the second term, the a priori probability $\rho(\bar{Z})$, incorporates information about the subsurface that is independent of the newly observed data from which the interferences are being made. In the event that the Gaussian model is an acceptable approximation for all uncertainties of the problem, the likelihood function is given by

$$L(\bar{Z}) \propto \exp\left(-\frac{1}{2}\{d_{obs}-g(\bar{Z})\}^T C_{dat}^{-1}\{d_{obs}-g(\bar{Z})\}\right) \quad (1.a)$$

$$\rho(\bar{Z}) = \exp\left(-\frac{1}{2}\{\bar{Z}-\bar{Z}_{prior}\}^T C_{mod}^{-1}\{\bar{Z}-\bar{Z}_{prior}\}\right) \quad (1.b)$$

where, $g(\mathbf{Z})$ is the electromagnetic forward modeling operator that generates the displacement field given a discretized model \bar{Z} of the subsurface, d_{obs} is the synthetic wave recorded by the antenna, C_{dat} is the data covariance matrix, \bar{Z}_{prior} is the center of the a priori probability density, and C_{mod} is the model covariance matrix.

The expected electric permittivity values to be obtained depend on the frequency content of the wavefield used to probe the electrical properties of the material, both its real and its imaginary part. The imaginary part of the electrical permittivity is negligible compared to the real part for frequencies of 10Hz and 100MHz. For this frequency range, the imaginary part is disregarded in materials where the displacement currents prevail with respect to the conduction currents, applicable to resistive materials, but in conductive materials the imaginary part is important and cannot be disregarded.¹ In this work we model clay, sand and pyroplastic materials, in which the displacement currents predominate and the Ricker wavelet is defined with a predominant frequency of 100MHz. In this way the imaginary part of the electric permittivity is neglected, taking into account only the real part of the electrical permittivity, that is, only the polarization component within the material is considered and the dielectric losses are disregarded.

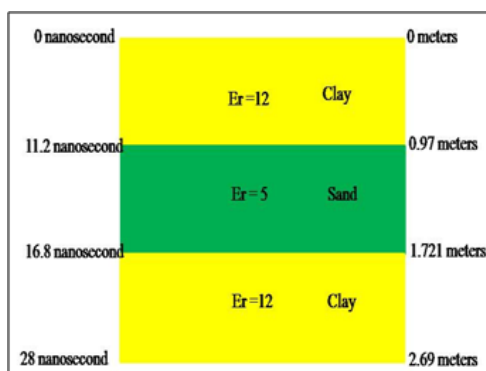


Figure 1 The parametrization of the physical model is shown. Due to the assumed lithology structure, three reflectors are present as follows:

1. The air-clay reflection located at 0ns;
2. The reflection of the clay-sand contact located at 11.2ns and;
3. The deeper reflection of the sand-clay contact located at 16.8ns.

Table 1 Relative electrical permittivity as a function of sample

Relative electrical permittivity	tm (sample)
$E_r = 12 + 0.1 * \sin(12 * 0.7 * tm)$	0-16
$E_r = 5 + 0.1 * \sin(12 * 0.7 * tm)$	17-24
$E_r = 12 + 0.1 * \sin(12 * 0.7 * tm)$	25-40

Geostatistical inversion

To find the electromagnetic independence, we have that the synthetic wave captured by the radio-receiver is d_{obs} :

$$d_{obs} = \begin{pmatrix} X[0] \\ X[1] \\ \vdots \\ X[n] \end{pmatrix} \quad (2)$$

Model: clay-sand-clay

This three-layer model has property material as clay loam at the top layer, with a conductivity of $\sigma=9$ MS/m and an average relative dielectric constant, $E_r=12$, which is 0.97m thick. In the middle layer there is a sand formation, $\sigma=5$ MS/m and $E_r=5$, with 0.75 m thickness, this formation makes contact at 1.72m depth with a half space characterized by a clay formation with the same electrical properties of the clay given at the top layer, see Figure 1 below. Below in Table 1, it is shown the distribution of relative electrical permittivity as a function of sample. Note that $(tm * 0.7)$ gives the time in nanoseconds. For this work the sampling frequency is $f_s=1.43$ cycles/nanosecond and the maximum frequency of the synthetic generated wavefield and picked-up by the antenna on the surface is $f_{max} = 0.6$ cycles/nanosecond, which complies with the Nyquist sampling theorem. There is an aliasing lower than 5% in the region of the frequency of 0.6 cycles per nanosecond, therefore it is considered optimal to comply with the sampling theorem of Nyquist, see Figure 2.

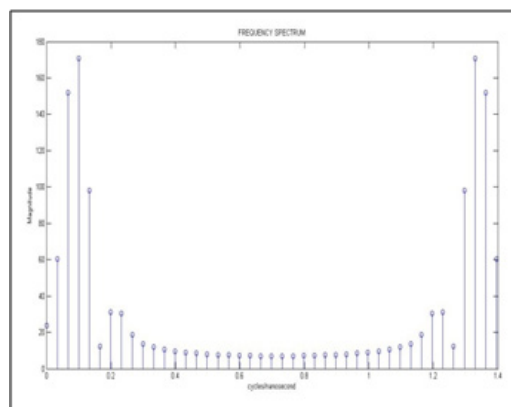


Figure 2 Shows the module of the fast Fourier transform versus the frequency in cycles/nanoseconds for synthetic electromagnetic wave without noise.

The synthetic “observed” wave data, at the study site, is approximated by the convolution model:

$$g(\vec{Z}) = \begin{pmatrix} Y[0] \\ Y[1] \\ \vdots \\ Y[n] \end{pmatrix} \quad (3)$$

The function to be optimized and that involves the synthetic “observed” wave data is:

$$S_1(\vec{Z}) = \frac{1}{2} \left\{ d_{obs} - g(\vec{Z}) \right\}^T C_{dat}^{-1} \left\{ d_{obs} - g(\vec{Z}) \right\} \quad (4)$$

The function to be optimized and related to the prior well information is given by:

$$S_2(\vec{Z}) = \frac{1}{2} \left\{ \vec{Z} - \vec{Z}_{previo} \right\}^T C_{mod}^{-1} \left\{ \vec{Z} - \vec{Z}_{previo} \right\} \quad (5)$$

The permittivity distribution within the well, after a size N window average, is

$$Z_{np} [k] = \frac{1}{N+1} \sum_{i=k-N/2}^{k+N/2} Z_p [i] \quad (6)$$

Where, “k” represents the specific number of the permittivity sample in the well.

For “k” = 1, (7)

$$Z_{np} [1] = \frac{1}{\left(\frac{N}{2}+1\right)} \sum_{i=1}^{1+N/2} Z_p [i] \quad (7)$$

Must be taken into account when $1 < k < N/2$, to what is the $Z_{np} [k]$.

For “k” = n,

$$Z_{np} [n] = \frac{1}{\left(\frac{N}{2}+1\right)} \sum_{i=n-N/2}^n Z_p [i] \quad (8)$$

Must be taken into account when $k > \text{length of } Z_p - N/2$, to what is the $Z_{np} [k]$.

Calculate the difference of the synthetic electrical permittivity data of well Z_p and the average Z_{np} :

$$L_p [k] = Z_p [k] - Z_{np} [k] \quad (9)$$

The covariance matrix for the model is defined as:

$$C_{mod} (i, j) = \frac{1}{n+1} \sum_{k=1}^n L_p [k] * L_p [k+j-i] \quad (10)$$

where n = number of the sample of the synthetic electrical permittivity data of well.

The inverse of the covariance matrix C_{mod} is defined as:

$$C_{mod}^{-1} (t, t') = \frac{1}{\sigma_{mod} (t, t')} \quad (11)$$

The inverse of the covariance matrix of electromagnetic wave data C_{dat} is defined as:

$$C_{dat}^{-1} (t, t') = \frac{1}{\sigma (t, t')} \quad (12)$$

In this work, the C_{dat} covariance matrix is diagonal, whose elements of the diagonal is taken equal to one, for this model, it can also be taken as the variance of the synthetic trace that would be recorded by the geophone.

The total function to be optimized is given below:

$$S(\bar{Z}) = S_1(\bar{Z}) + S_2(\bar{Z}) \quad (13)$$

Developing equation (4) by indicial product and replacing equations (2) and (3) we have

$$\begin{aligned} S_1(\bar{Z}) &= \frac{1}{2} \Delta X^T C_{dat}^{-1} \Delta X \\ &= \frac{1}{2} \Delta X_i e_i^T \cdot (C_{dat}^{-1})_{ls} e_l e_s^T \cdot \Delta X_j e_j \\ &= \frac{1}{2} \Delta X_i \delta_{il} (C_{dat}^{-1})_{ls} \Delta X_j \delta_{sj} \\ &= \frac{1}{2} \Delta X_i (C_{dat}^{-1})_{ij} \Delta X_j \\ &= \frac{1}{2} \sum_{i=0}^n \sum_{j=0}^n \Delta X_i (C_{dat}^{-1})_{ij} \Delta X_j \\ &= \frac{1}{2} \sum_{i=0}^n \sum_{j=0}^n (X[i] - Y[i]) (X[j] - Y[j]) \frac{1}{\sigma(i, j)} \\ &= \frac{1}{2} \sum_{t'=0}^n \sum_{t=0}^n \frac{(X[t'] - X[t]) (X[t] - Y[t]) - (X[t] - Y[t]) (X[t'] - Y[t'])}{\sigma(t', t)} \end{aligned} \quad (14)$$

We take the gradient of the function (14), “S1” with respect to $Z[m]$ and replacing the following equation

$$Y[t] = \sum_{\tau=0}^n W[t-\tau] \frac{(Z[\tau+1] - Z[\tau])}{(Z[\tau+1] + Z[\tau])} \quad \text{before of take the gradient of}$$

the function (which is the convolution of the source wavelet with the synthetic electrical permittivity data which you want to find) is:

$$\begin{aligned} \frac{\partial S_1(\bar{Z})}{\partial Z[m]} &= -\frac{1}{2} \sum_{t'=0}^n \sum_{t=0}^n \frac{X[t'] W[t-\tau]}{\sigma(t', t)} \left\{ (\delta_{m, \tau+1} - \delta_{m, \tau}) (Z[\tau+1] - Z[\tau])^{-1} - (Z[\tau+1] - Z[\tau]) (Z[\tau+1] + Z[\tau])^{-2} \right. \\ &\quad \left. (\delta_{m, \tau+1} + \delta_{m, \tau}) \right\} - \frac{1}{2} \sum_{t'=0}^n \sum_{t=0}^n \frac{X[t] W[t'-\tau']}{\sigma^2(t', t)} \left\{ (\delta_{m, \tau'+1} - \delta_{m, \tau'}) (Z[\tau'+1] + Z[\tau'])^{-1} - (Z[\tau'+1] - Z[\tau']) (Z[\tau'+1] + Z[\tau'])^{-2} \right. \\ &\quad \left. (\delta_{m, \tau'+1} + \delta_{m, \tau'}) \right\} \\ &\quad + \frac{1}{2} \sum_{t'=0}^n \sum_{t=0}^n \sum_{\tau=0}^n \frac{W[t'-\tau'] W[t-\tau]}{\sigma(t', t)} \left\{ (Z[\tau+1] - Z[\tau]) (Z[\tau+1] + Z[\tau])^{-1} [(\delta_{m, \tau'+1} - \delta_{m, \tau'}) (Z[\tau'+1] + Z[\tau'])^{-1} - (Z[\tau'+1] - Z[\tau']) (Z[\tau'+1] + Z[\tau'])^{-2}] \right. \\ &\quad \left. (\delta_{m, \tau'+1} + \delta_{m, \tau'}) \right\} + (Z[\tau'+1] - Z[\tau']) (Z[\tau'+1] + Z[\tau'])^{-1} [(\delta_{m, \tau'+1} - \delta_{m, \tau'}) \\ &\quad - (Z[\tau+1] - Z[\tau]) (Z[\tau+1] + Z[\tau])^{-2} (\delta_{m, \tau'+1} + \delta_{m, \tau'})] \end{aligned} \quad (15)$$

The gradient of the equation (5) is:

$$\frac{\partial S_2}{\partial Z[m]} = \frac{1}{2} \sum_{t=0}^n (z[t] - z_p[t]) \left(\frac{1}{\sigma_{mod(m,t)}} + \frac{1}{\sigma_{mod(t,m)}} \right) \quad (16)$$

To optimize the equation (13) the method of the Escalated Conjugate Gradient is used, having that the gradient

of the equation (13) is the sum of the equations (15) and (16).

Results

To test our computer code but it up with expressions given above, we used the Ricker source wave function with a predominant frequency of 0.1 cycles/nanosecond, and is shown in Figure 3. The synthetic wave that simulates the wave recorded in the surface for the acquisition system (synthetic electromagnetic wave without noise) is shown in Figure 4, and is obtained from the convolution of the Ricker wave with the relative electrical permittivity of the geological model (Figure 1). In Figure 5 we present both the relative electric permittivity model (perturbation from the initial reference model) of the well, and the one obtained in a second step by means of a moving window average with 2 samples. These models were used to estimate the covariance matrix of the model, through the equations from (6) to (10). The calculated relative electrical permittivity data is made through the optimization method of the scaled conjugate gradient,² and the permittivity was initialized at a constant value equal to the average of the relative synthetic electrical permittivity data of well. The Figure 6 shows the initialization data of the optimization algorithm (light blue), the synthetic data of real permittivity (line dark green, real synthetic data), the wellbore permittivity data averaged with the moving window (line red), and the calculated permittivity data (line dark blue)

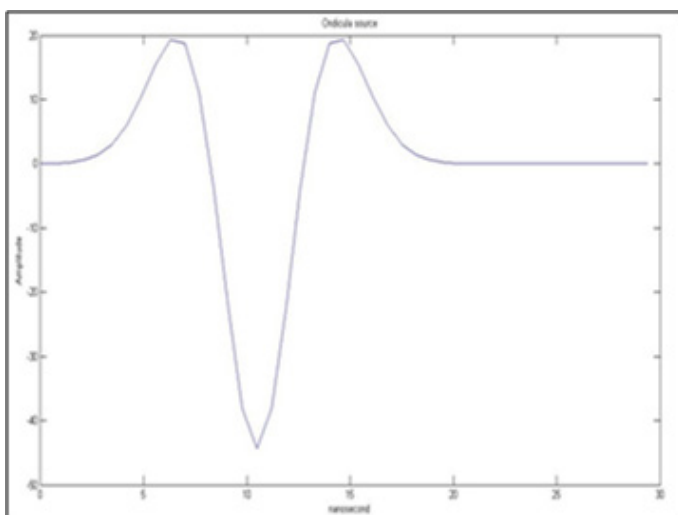


Figure 3 Source time function in terms of Ricker pulse, given in nanoseconds.

The initialization data of the optimization algorithm has a constant value of relative permittivity of 11.2, which is observed in Figure 7, and in which the red line shows the well permittivity data averaged with the moving window, and in color dark blue represents the calculated permittivity data for the synthetic electromagnetic wave without noise. When the program iterates, the initial data changes until the calculated data is obtained. In figures 8 and 9, we show the real and calculated synthetic relative electrical permittivity, respectively, as also is observed for the geological model (Figure 1). Only 5 electric permittivity inversions were calculated and 38 equal to the previous ones were placed continuously (it is a means of parallel layers) to make Figures 8-16 shown below. The synthetic wave that simulates the wave recorded at the top surface (synthetic electromagnetic wave with noise) is shown in Figure 10, and is obtained from the convolution of the Ricker wave with the relative electrical permittivity of the geological model (Figure 1), plus random Gaussian noise of amplitude 5% of the maximum amplitude of the synthetic wave without noise. For this model the sampling frequency is $f_s=1.43$ cycles/nanosecond and the maximum frequency of the synthetic wave (with noise) recorded on the surface is $f_{max}=0.6$ cycles/nanosecond, which complies with the Nyquist sampling theorem. There is an aliasing lower than 5% in the region of the frequency of 0.6 cycles per nanosecond, therefore it is considered optimal to comply with the sampling theorem of Nyquist, see Figure 11.

The stopping criterion is taken when the gradient of the function target is minor that the number of observations (variables), also can be when:

$$\left\{ d_{obs} - g(\bar{Z}) \right\}^T C_{dat}^{-1} \left\{ d_{obs} - g(\bar{Z}) \right\} \leq N \quad (17)$$

Where N is the number of observations.

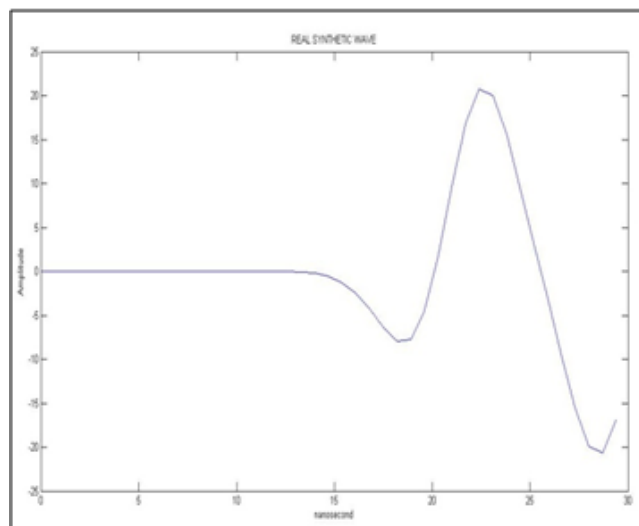


Figure 4 Synthetic "recorded" wave form (without noise) versus time in nanoseconds.

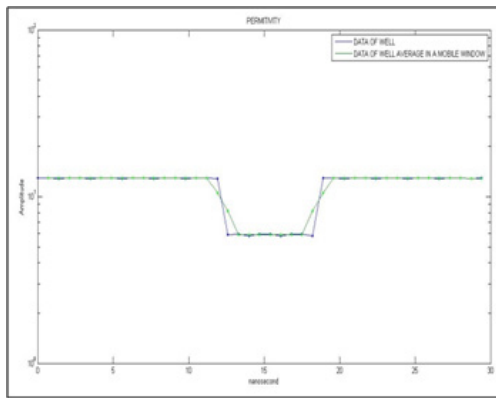


Figure 5 Relative electric permittivity in the well and its average made with a moving window that uses 2 samples, these are a function of nanosecond time.

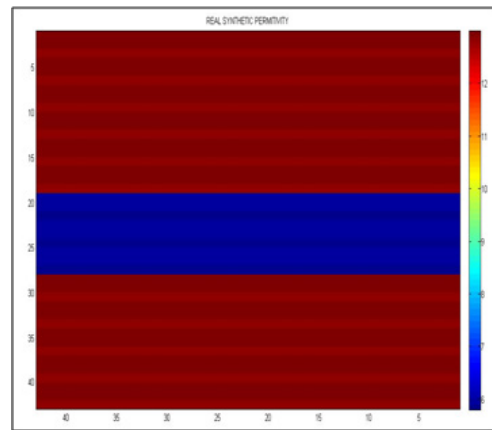


Figure 8 It shows the real synthetic permittivity, the left vertical axis must be multiplied by 0.7 for that of the scale in nanoseconds, and the horizontal axis is separated by centimeters.

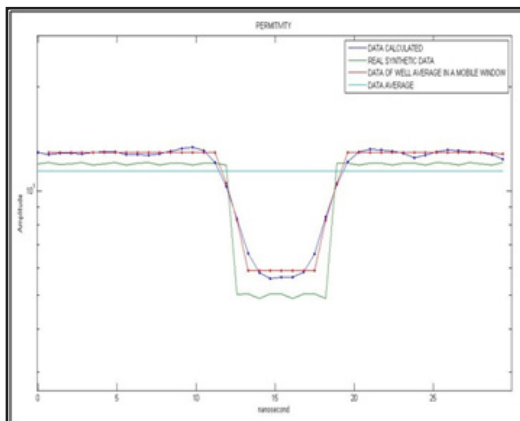


Figure 6 Shows the initialization data of the optimization algorithm (light blue), the synthetic data of permittivity of the model (dark green, real synthetic data), the averaged relative electric permittivity of well in a mobile window (red), and the calculated permittivity data (dark blue) for the synthetic electromagnetic wave without noise, all as a function of nanosecond time.

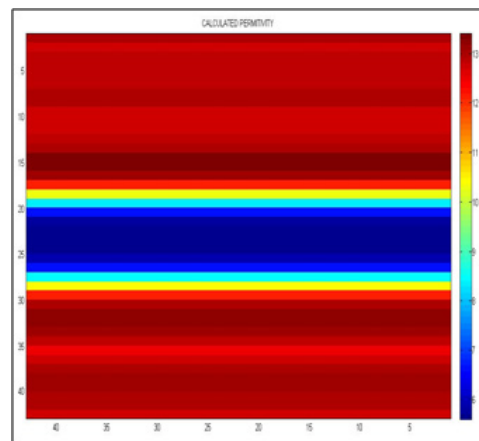


Figure 9 It shows the calculated permittivity, the left vertical axis must be multiplied by 0.7 for that of the scale in nanoseconds, and the horizontal axis is separated by centimeters. We can see the interfaces between clay-sand at 12 nanosecond and sand-clay at 18 nanosecond.

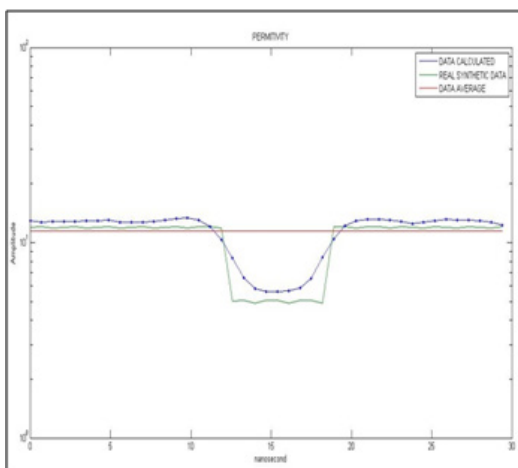


Figure 7 Shows the calculated permittivity data (dark blue line) for the synthetic electromagnetic wave without noise, real synthetic permittivity (dark green line) and average permittivity (the initialization data of the optimization algorithm; red line). The scale of the vertical axis is in logarithmic.

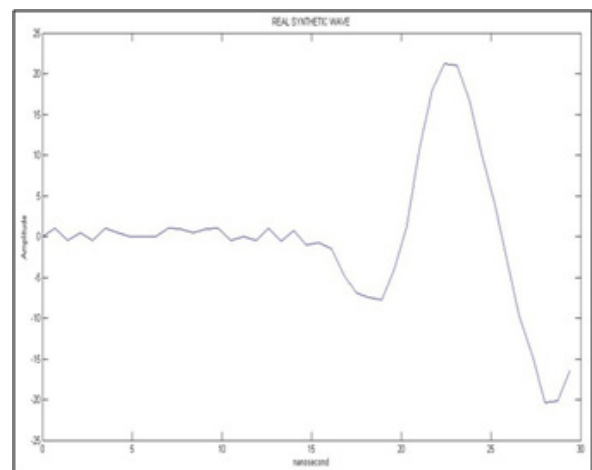


Figure 10 Shows the synthetic wave recorded by the geophone (with noise) versus time in nanoseconds.

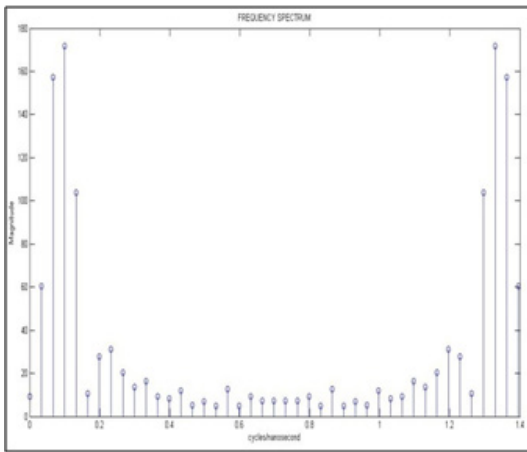


Figure 11 Shows the module of the fast Fourier transform versus the frequency in cycles/nanoseconds for synthetic wave with noise.

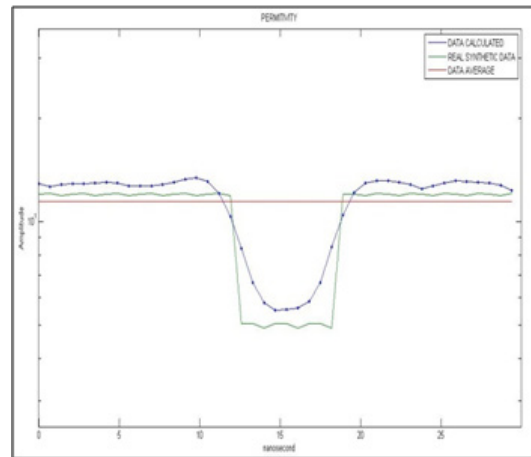


Figure 14 Shows the calculated permittivity data (dark blue line) for the synthetic electromagnetic wave with noise, real synthetic permittivity (dark green line) and average permittivity (red line). The scale of the vertical axis is in logarithmic.

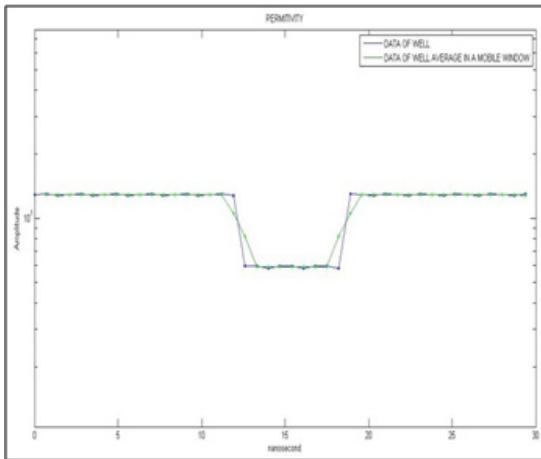


Figure 12 Shows the relative electric permittivity of the well and its average made with a mobile window that uses 2 samples, these are a function of nanosecond time. The scale of the vertical axis is in logarithmic.

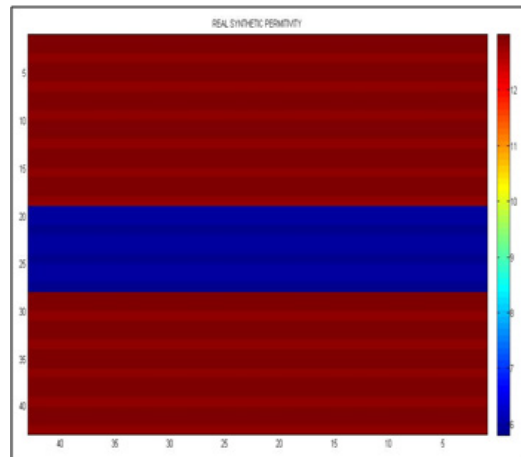


Figure 15 It shows the real synthetic permittivity, the left vertical axis must be multiplied by 0.7 for that of the scale in nanoseconds, and the horizontal axis is separated by centimeters.

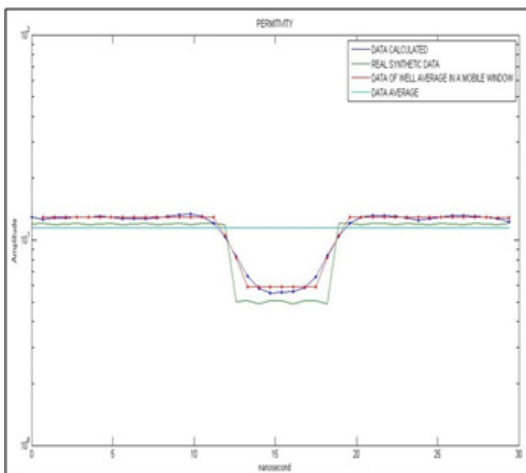


Figure 13 Shows the initialization data of the optimization algorithm (light blue), the synthetic data of real permittivity (dark green), the electric permittivity data averaged in a moving window (red), and the calculated permittivity data (dark blue) for the synthetic electromagnetic wave with noise, all as a function of nanosecond time. The scale of the vertical axis is in logarithmic.

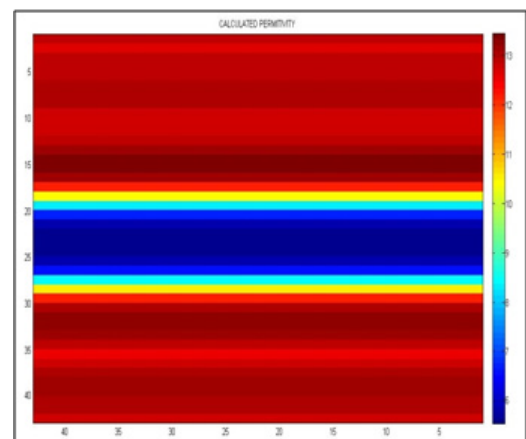


Figure 16 It shows the calculated permittivity, the left vertical axis must be multiplied by 0.7 for that of the scale in nanoseconds, and the horizontal axis is separated by centimeters. We can see the interfaces between clay-sand and sand-clay at 12 nanosecond and sand-clay at 18.5 nanosecond approximately.

Conclusion

The covariance matrix of the synthetic data is diagonal and its variance was chosen equal to one, while the covariance matrix of the model has a value of 0.5 on its diagonal, the latter being smaller which causes the permittivity calculated for the Synthetic wave with and without noise, fit fairly close to the electric permittivity of the averaged well produced by the moving window (this averaging is the center of the a priori probability density), and give acceptably close to the real synthetic permittivity. It is observed that the scaled conjugate gradient method works well for this type of optimization. It is observed that with this inversion technique the medium can be characterized, and if the complex component of the permittivity is calculated by some mathematical model, the characterization of the medium would be extended for other types of soil and frequencies. The RAM memory of the laptop is 2 GHz, and Intel Core Duo, the run took 170 minutes per trace. Generally speaking, higher performance requires more memory and more running time. In the future, the issue of running time can be further addressed by the distributed parallel algorithm or the GPU implementation.

Acknowledgements

None.

Conflict of interests

The author declares there is no conflict of interest.

References

1. Daniels D, Gunton D, Scott H. Introduction to subsurface radar. *IEE Proceedings*. 1998;135(4):278–320.
2. Martin F Moller. A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. *University of Aarhus, Neural Networks*. 1993;6(4):525–533.