

Research Article

A Review of Soft Techniques for Electromagnetic Assessment of Concrete Condition

**F. A. A. Queiroz,¹ D. A. G. Vieira,¹ X. L. Travassos,²
and M. F. Pantoja³**

¹ Parque Tecnológico de Belo Horizonte, Rua Professor José Vieira de Mendonça N 770, Salas 406 e 407, Engenho Nogueira, 31310-260 Belo Horizonte, MG, Brazil

² Departamento Regional da Bahia, Centro Integrado de Manufatura e Tecnologia, SENAI, Avenida Orlando Gomes N 1845, Piatã, 41650-010 Salvador, BA, Brazil

³ Departamento de Electromagnetismo y Física de la Materia Facultad de Ciencias, Universidad de Granada Campus de Fuentenueva, Severo Ochoa s/n, 18071 Granada, Spain

Correspondence should be addressed to F. A. A. Queiroz, francisco.queiroz@enacom.com.br

Received 26 July 2012; Revised 5 November 2012; Accepted 14 November 2012

Academic Editor: Fatih Yaman

Copyright © 2012 F. A. A. Queiroz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper describes some efforts to design more efficient analysis for the GPR raw-data interpretation. Such analysis requires algorithms by which problems having complex scattering properties can be solved as accurately and as quickly as possible. This specification is difficult to achieve when dealing with iteratively solved algorithms characterized by a forward solver as part of the loop, which often makes the solution process computationally prohibitive for large problems. The inverse problem is solved by an alternative approach which uses model-free methods based on example data. Measurements with GPR equipment were performed to validate the algorithms using concrete slabs.

1. Introduction

Concrete is a highly versatile construction material well suited for many applications. Strength, durability and many other factors depend on the relative amounts and properties of the individual components. It is almost impossible to predict its final electrical characteristics at the job site. It is electrically nonconductive but usually contains significant amounts of steel reinforcement. While concrete is noted for its durability, it is susceptible to a range of environmental degradation factors, which can limit its service life.

Throughout the literature, radar technology proved to be the cheapest and easiest method for mapping reinforcements but neither characterization of flaws by dimension and material nor crack detection could be demonstrated [1–5]. The concept of using Ground

Penetrating Radar (GPR) for the nondestructive testing (NDT) of concrete structures has grown in the past years, involving researchers from many different disciplines. The interest in this technique can be explained, basically, by its advantages compared to other techniques such as portability, relative low cost and weight, the ease of application to different surfaces. However, the imaging of typical field data may be difficult due to problems like limited coverage, noisy data, or nonlinear relations between observed quantities and physical parameters to be reconstructed. Several algorithms for reconstructing an unknown object have been implemented, most based on Newton-type iteration schemes what requires significant computational burden and a high level of complexity.

Nevertheless, radar has significant potential for development by way of software for signal and image processing to improve resolution solving an inverse problem.

The inverse problem consists in using the results of actual observations to infer the values of the parameters characterizing the system under investigation. Estimating inclusions characteristics using measured signals obtained by scanning antennas is known as an ill-posed inverse problem. Electromagnetic wave propagation inverse problems are typically ill-posed, as opposed to the well-posed problems more typical when modeling physical situations where the model parameters or material properties are known.

Some different reference techniques of non-destructive evaluation of reinforced concrete structures were collected in the review book [6]. One of those works is about migration algorithms [7], which can be used in simulated and real GPR data to focus the information contained in hyperbola, facilitating the visual location and assessment of inclusions in concrete.

Kirchhoff migration and Synthetic Aperture Focusing Technique (SAFT) [8] help to identify inclusions in the host medium raw data. The latter method tries to enhance the spatial resolution of inclusions whereas the first method concentrates in one point the consecutive hyperbolic echoes of a cylindrical inclusion.

For example, in [9, 10], a migration algorithm was used to define the number of inclusions and locate their centers inside nonhomogeneous host medium. Although that type of medium may affect the results of these algorithms, the obtained information was then used in the subsequent phase wherein the geometry and the electrical characteristics of inclusions were estimated by other method.

The general problem of testing concrete using subsurface radar including assisting with intelligent systems is addressed in [11]. Emphasizing the interest in developing methods to solve one of these inverse problems, many works including Computational Intelligence were reviewed in [12] with the specific objective of rebar detection in concrete using GPR, which is the same topic treated in [13] using Neural Networks. A Multilayer Perceptron model was used in [14] as a simplified hyperbolic shape detector in GPR images to identify steel reinforcing bars in concrete. As plastic pipes can be found inside concrete structures too, their identification in GPR images is a similar problem whose solution is proposed in [15] using multi-agent systems.

In [16], three works using Neural Networks in GPR data sets to inspect highways, bridge decks, and pavements are listed and can find applications in these structures sometimes made of concrete. This is the case considered in [17] to be solved by using Independent Component Analysis and fractals to detect defects in bridge decks. A Fuzzy System was developed in [18] to merge data from GPR and infrared and digital inspection system in order to detect problems in mountain tunnel lining inspected. In [19], water and chloride contents were predicted by Neural Networks from radar data to evaluate physical

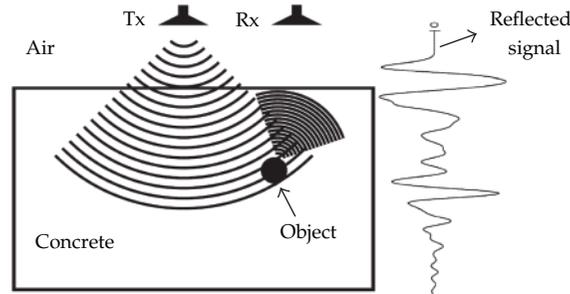


Figure 1: GPR inverse problem description.

condition of concrete. Thus, much effort has been done to improve automated concrete inspection by developing solutions for this inverse problem.

The aim of this paper is to expose efficient inverse methods to GPR raw data of concrete. The analysis steps of raw radargrams involve a priori information, subjective criteria, and indefinite sequences of data processing steps due to specific set of scan conditions. That makes it hard to completely characterize the inclusions in concrete using simple models. Therefore, this work describes alternative ways of solving the inverse problem of retrieving buried inclusions in dielectrics with electrical characteristics similar to concrete slabs.

The information extracted from the inverse models is used to estimate the electrical parameters of the probed materials inside concrete. It may help the development of software to aid analysis of GPR data from non-destructive evaluation of concrete structures.

The inverse problem considering the GPR assessment of concrete is outlined. The concept of inverse problem is defined with respect to the inclusion reconstruction in concrete slabs. Three reconstruction algorithms are implemented and experimental tests are performed to investigate the performance of the proposed algorithms.

2. Problem Settings

The objective of this work is the inverse problem of finding the characteristics of a target buried in a dielectric material given the reflected field measured by the antenna (Figure 1). In practice, the reflected signal is a collection of discrete observations in time. For the radar assessment of concrete, the objective is to determine a finite number of parameters. The parameters needed to characterize inclusions in a dielectric slab are found by identifying electrical (permittivity and conductivity) and geometrical (depth and radius) properties.

Such a problem is called discrete inverse problem or parameter estimation problem. In general, this is a difficult problem because the obtained information is not sufficient for estimation, which requires some a priori information about the inclusions. On the other hand, a large number of data can bring the problem to instability.

One possible technique to overcome these shortcomings is the use of parametric algorithms which are based on optimization algorithms. This kind of algorithms update the parameters iteratively to minimize (or maximize) a certain evaluation function. The computation of an approximate electric or magnetic field is done by the finite-difference time-domain (FDTD) method as a forward solver.

The optimization process can be carried out by using a variety of algorithms such as the Newton method, conjugate gradient method [20], or evolutionary algorithms. One major

limitation of using parametric algorithms is the time of calculation that can be prohibitive for 3D problems. Nonparametric algorithms can also be used. Usually, they are complex to implement but they can solve inverse problems faster than parametric algorithms because it is not necessary to iteratively evaluate the function.

Those approaches are less efficient in the sense of using all the available raw data without extract relevant prior information. So, they do not contribute to the development of more efficient models that are closer to white boxes. An alternative way is to use feature selection algorithms, such as Simba and Relief [21], to select characteristics from a scattered GPR wave. These are based on the concept of margin to define the relevant features of the data.

3. Parameter Estimation Approaches

3.1. Model Fitting

Model fitting method is one of the parametric imaging algorithms which utilize the evaluation function and optimization algorithms. In the model fitting method, target characteristics and location are expressed as parameters. The parameters are updated iteratively to minimize the difference between the observed data and the estimated data.

An incident wave and a scattered wave can be used to characterize the scattering object. Usually, in real world problems, the incident and scattered waves are known and it is desired to identify the scattering object. This problem can be written as an optimization problem involving the scattered wave of the unknown object $E(\theta_0)$, the reference object, and the scattered wave of a test object $E(\theta)$ [9].

Thus, θ^* , the optimum θ , is the argument that minimizes the error of the reference object scattered wave $E(\theta_0)$ relative to the test object scattered wave $E(\theta)$. Mathematically:

$$\theta^* = \arg \min f(\theta) = \sum_{i=1}^{ns} (E(\theta_0) - E(\theta))^2, \quad (3.1)$$

where ns is the number of sample points where the scattered wave is measured. Note that $E(\theta_0)$ is known even though θ_0 is unknown. The scattered field $E(\theta)$ is then generated assuming one tests θ , and the optimization procedure aims at minimizing the error between $E(\theta_0)$ and $E(\theta)$ in such a way as to identify the scatter object θ_0 .

The problem described in (3.1) is usually multimodal, as shown in Figure 2, where the unknowns are the radii of two inclusions given their physical properties. This multi-modal characteristic motivates the use of a stochastic approach instead of a deterministic one. This problem was solved in this paper using the Particle Swarm Optimization (PSO), which is described next.

The Particle Swarm Optimization (PSO) algorithm is similar to genetic algorithms (GAs) due to the random initialization. The first difference is that each potential solution is called particles, instead of individuals, and they “fly” on the search space. To each particle of the swarm during the iterations, the position of the best solution found to a given particle, called *pbest* (particle best), is saved.

The best value found considering all the particle is also saved, and is called *gbest* (global best). At each iteration, the PSO is based on the change in the particle’s velocity in the direction of its *pbest* and *gbest*, weighted by random terms.

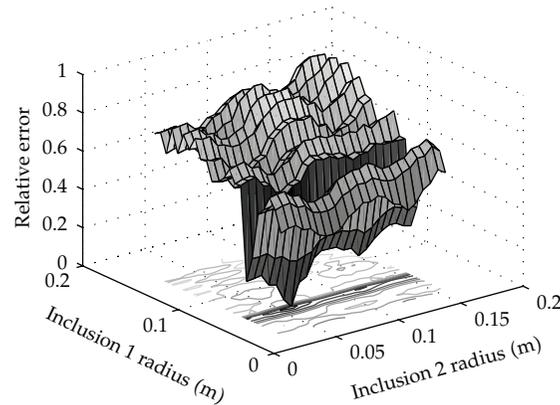


Figure 2: Surface plot of (3.1) representing the problem of finding the radii of two inclusions, given that their location and physical properties are known. Even in this simple example the surface is clearly multi-modal, thus, stochastic methods must be employed in the parametric approach.

Table 1: Inverse scattering problem definitions.

| | Depth (mm) | Radii (mm) |
|-------------|------------|------------|
| Min | 50 | 25 |
| Max | 250 | 100 |
| Ref. Object | 100 | 45 |

In our radar problem, the unknown object was obtained by experiment at depth of the inclusion equal to 100 mm, the radius equal to 45 mm and the standard deviation of the nonhomogeneous medium, sd , equal to 0.15. The target considered was water. The definition of the reference object as well as the range of the variables for the optimization process are summarized in Table 1.

The PSO was initialized using 50 particles over 50 iterations. The *gbest* evolution throughout the iterations is shown in Table 2. In the 50th iteration, the algorithm was capable of finding a solution very close to the desired one.

Nonetheless, good approximations were already available in the 5th iteration; thus, if a very accurate outcome is not necessary the algorithm would need only about 5 iterations to converge. It was studied in [10] an algorithm to mix migration algorithms with PSO in order to minimize the computational burden of this kind of investigation.

Although this methodology for model fitting may get better results by exploring a wide range of values for the independent parameters, better results are not guaranteed and the computer resources (hardware and time) needed to optimize the problem may not be available.

3.2. Artificial Neural Networks

Several imaging or inversion techniques have been developed to refocus the scattered signals back to their true spatial location and then decrease the interpretation time for effective maintenance and/or repair. Among them the Artificial Neural Networks (ANN) has been

Table 2: PSO results.

| Iteration | Depth (mm) | Radii (mm) |
|-------------|------------|------------|
| 1 | 87 | 32 |
| 5 | 97 | 44 |
| 10 | 97 | 47 |
| 50 | 100 | 44.4 |
| Ref. Object | 100 | 45 |

proven to be a promising technique in the solution of inverse electromagnetic field problems [22].

In this section it is addressed a 2D problem where a cylinder of unknown characteristics is buried in a nonhomogeneous dielectric. The incident and scattered wave are simulated using FDTD to train the ANN. The dielectric medium uses the electrical characteristics of concrete [1] with a mean relative electrical permittivity value of 6 and standard deviation 0.15, that is, a non-homogenous medium [10].

The investigation domain is illuminated by a differentiated Gaussian pulse with a center frequency of $f = 900$ MHz and bandwidth between 0.3 and 2 GHz. In order to control the numerical dispersion and provide good discretization for the inclusions the spatial steps were chosen as $\Delta x = \Delta y = 6$ mm. The aim of this problem is, given an incident wave, W_i , and scattered wave, W_s , to determine the radius and depth of the inclusion.

To train the ANN a set of different inclusion examples, say $S(\cdot)$, for different radii, depths, ϵ_r , and σ were generated.

The ANN has been trained with a set of different inclusion examples, constructed by varying the radius in the range [0.02–0.1] m according to the rule $\text{radius} = 0.02 + i \times 0.001$, $i = 0, \dots, 80$, with ϵ_r in the range [1–10], according to the rule $\epsilon_r = 1 + i \times 1$, $i = 0, \dots, 9$, σ in the range [0–4000] S/m according to the rule $\sigma = 0 + i \times 500$, $i = 0, \dots, 8$ and depth in the range [0.05–0.25] m according to the rule $\text{depth} = 0.05 + i \times 0.025$, $i = 0, \dots, 9$ summing a total of 1640 examples.

The W_s is the only information available in the real cases; therefore, it has to be used to characterize the inclusion. In this paper 1200 time steps were considered; thus, a problem in \mathbb{R}^{1200} must be solved. The ANNs suffer from a phenomenon called the curse of dimensionality [23], that is, the learning process becomes slower and less effective.

The dimensions of data features have great impact in samples density in input dimension and in the number of parameters in the structure of a chosen model. So, the samples with more features get more distant from each other and even if numerous samples are available in information extraction process, the number of features may require great computer resources to represent and compute model structure during parameters tuning in training.

To reduce the dimensionality of the problem Principal Component Analysis (PCA) was used. The main advantage of PCA is that once these patterns in the data are found, the data can be compressed, that is, by reducing the number of dimensions, without much loss of information. This technique is commonly used in image compression.

The input vectors are first normalized so that they have zero mean and unity variance. For PCA to work properly, one has to subtract the mean from each of the data dimensions. The PCA uses a linear mapping of a given set of samples $S_q = \{x_1, \dots, x_T \mid x_i \in \mathbb{R}^p\}$ to construct a new data set $S_p = \{y_1, \dots, y_T \mid y_i \in \mathbb{R}^q\}$, where $q \leq p$.

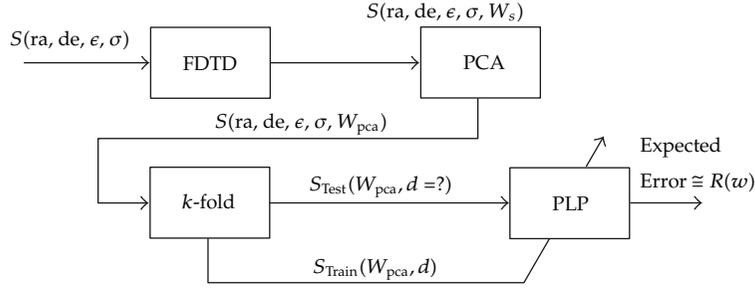


Figure 3: This figure shows an overview of the detection system employed using the ANN. Firstly, given a set S of radii (ra), depths (de), ϵ_r , and σ a scattered wave ($W_s \in \mathbb{R}^p$) is calculated using FDTD. Afterwards, a dimensional reduction is applied. In the next step, the expected error is estimated to adapt the Parallel Layer Perceptron network parameters. In this work a 10-fold cross-validation was employed.

Another interpretation of the PCA is the construction of directions that maximize the variance. The transformation V_q generates a projection space in which the covariance matrix is diagonal. The diagonal covariance matrix implies that the variance of a variable with itself is maximized while it is minimized with respect to any other variable. Thus, the q variables with higher variance in the new space should be kept. The principal components of a set of data in \mathbb{R}^p provide a sequence of best linear approximations to that data, of all ranks $q \leq p$.

The problem of dimension reduction and ANN training has its scheme shown in Figure 3. Considering the input space initially in \mathbb{R}^{1200} , it can be projected in a \mathbb{R}^{286} without any loss of information, that is, 100% of the data variance was kept. Considering 99.99% of the variance, the variables can be projected in \mathbb{R}^{139} and in \mathbb{R}^{51} when 99% of the original variance is kept. These are remarkable reductions that help in reducing the curse of dimensionality.

To evaluate the performance of the studied techniques the following error (loss) figure is used:

$$L(d_r) = \frac{|d_t - d_r|}{d_r}, \quad (3.2)$$

where d is the unknown variable (depth or radius), the subscript t indicates the real value of the variable, and the subscript r indicates the value reconstructed by the neural network. This measures the percentage deviation of the reconstructed object from the real one (desired object).

In Table 4, it is presented the results considering the Relative Error in the prediction of the depth given 286 dimensions, varying the number of neurons. It is clear that the error does not vary much as the number of neurons changes. The optimum is achieved for 9 neurons. The network starts overfitting afterwards. The number of neurons was determined using cross-validation. However, the results given different number of neurons are also acceptable results, which can be used in a real-world problem.

The ANNs have the advantages of having a numerous class of models, reaching high levels of accuracy due to their capacity of generalization, having universal approximation models for regression and classification models, can be fast computed due to their natural parallel structure (real time applications), being tolerant to structure failures (knowledge distributed in structure), having well known strategies to control their complexity,

having satisfactory models easy to understand and implement even for non-experts, may be retrained or modified once having new samples.

On the other hand, the drawbacks of ANNs are the necessity of having many representative pre-processed (normalized) samples to guarantee satisfactory understanding of domains of problems being modeled (the accuracy is conditioned to samples representing the problem), cannot always be easily understood or translated to a set of symbolic rules or equations due to their structure, cannot easily deal with many different problems, inputs, and outputs with different codification, noise levels, and approximation and from different areas, the learning process may need great computational resources and time.

3.3. Feature Selection

Feature selection (also known as subset selection) is a process commonly used in machine learning, wherein a subset of the features available from the data is selected for application of a learning algorithm. The best subset contains the least number of dimensions that most contribute to accuracy, discarding the remaining, unimportant dimensions. This is an important stage of pre-processing and is one of two ways of avoiding the curse of dimensionality (the other is feature extraction). Feature selection is, therefore, the task of choosing a small subset of features which is sufficiently to predict the target labels. This step is fundamental to build reliable classifiers.

Another phenomenon that takes place in high dimensional problems is the rise of the computational effort. Suppose that one wants to approximate the scattered wave for the 1640 samples. The numerous methods for modeling the nonlinear behavior are classically divided into three categories: physical, empirical, and table-based ones. Some may be difficult to categorize in this way, however, and therefore models are divided here simply into physical and empirical ones, and into black-box and circuit-level ones.

Feature selection aims at choosing for a given data set a subset which can capture the relevant information. The choice of features is important to avoid the curse of dimensionality and, therefore, guarantee a good convergence of the learning machines. It can also provide some understanding concerning the nature of the problem, as it indicates the main physical properties to classify an underground target.

In literature three types of Feature Selection algorithms are distinguished, the so-called wrapper, filter, and embedded methods [21–23]. Wrapper methods estimate the usefulness of a subset of features by a given predictor or learning machine. The Wrapper methods try to directly optimize the performance of a given predictor, this is done estimating the generalization performance. Even though it sounds a natural approach it can be unfeasible given its computational burden, as many classifier must be cross-validate.

Filter or Variable Ranking methods compute relevance scores for each single feature and choose the most relevant ones according to those scores. This can be done using usually ad hoc, evaluation function aiming at searching the set of features that maximize the information. Some of the commonly used evaluation functions are: the mutual information, the margin, dependence measures, among others.

The main drawback of such simple filter methods is that they are not able to detect inter-feature-dependencies; one example of inter-feature-dependency in our GPR problem is relation between the clutter and reflected echoes from near targets that can occur at the same time. Neither the first nor the second dimension alone helps to determine from which class an example is stemming, only both dimensions together contain enough information about the class membership.

Another drawback of those methods is its computational deficiency. But there are methods, which may be categorized as variable ranking methods and are also able to reveal such feature dependencies; for example, Relief [21] is one of those. This paper will consider feature selection algorithms based on the filter model. In this case the feature selection is a type of pre-processing, using some predefined cost function, as the classes separation margin.

The algorithms considered here are Relief and Simba as in [24]. The Relief algorithm is based on the statistical relevance of the features while Simba is based on the concept of the separation hypothesis-margin between two classes.

From the reflected wave, 15 features were computed to be selected.

- (i) Delay of the first reflection ($Delay$) [25],
- (ii) maximum amplitude ($\max(W_s)$) [25],
- (iii) reflected wave mean ($\text{mean}(W_s)$),
- (iv) reflected wave standard deviation ($\text{std}(W_s)$),
- (v) mean of the wave derivative ($\text{mean}(dW_s)$),
- (vi) standard deviation of the wave derivative ($\text{std}(dW_s)$),
- (vii) maximum amplitude of the Fourier transform (mFFT),
- (viii) energy in six different bands (B_1, B_2, \dots, B_6) of the Fourier transform,
- (ix) frequency of maximum amplitude of the Fourier transform (fmFFT),
- (x) signal energy ($\text{sen}(W_s)$).

The six most selected features in [26] were: $\text{mean}(W_s)$, mFFT, fmFFT, $\text{sen}(W_s)$, $\text{std}(dW_s)$, and $Delay$.

As already explained, the feature reduction avoids large structures to data set modeling in both regression and classification problems. In addition, a smaller set of features may expose models to less noisy data, and hence, it may improve classification accuracy. Other advantage is the reduction of computational resources needed to store data and calculated metrics from samples. And models with fewer features tend to facilitate obtaining more explainable models because the resultant features can be justified by known information which relates samples and their labels.

To do so, the feature selection methods must have the ability to deal with large data sets with many samples and features. Depending on the selected strategy to feature extraction, which replaces the original data by some of their metrics for evaluation, the new representation of each sample should not vary so much due to noise. But if it is not the case, the feature selection process may be affected by metrics intolerance to noise. The final subset of features may greatly vary besides requiring a consensus response and a more complex data pre-processing step. So, the feature selection methods must be applied more than once. Another problem of filter methods for feature selection is the requirement of having all classes well represented in training set. As the methods addressed in this work are based on hypothesis-margin, the samples near the margin should be presented in the set considered to rank features. The metric and strategy to compare the combinations of features have considerable impact on the quality of results and the number of evaluations. As a consequence, different filter methods can reach distinct results requiring almost similar time for evaluation. Then, by the problems exposed above, the final selected features are not guaranteed to be the best subset.

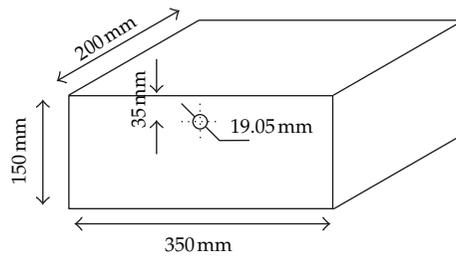


Figure 4: Concrete slab dimensions.

4. Real Data Processing

The concrete has been the most commonly used building material because it is durable, cheap, and readily available and can be cast into almost any shape. The main components of concrete are the hydrated cement paste, aggregates, water, and the transition zone.

The main constituent of concrete is cement, which is made up of calcium, silica, alumina, and iron oxide. The aggregates occupy 60 to 80% of the concrete's volume. They are usually stronger than cement and therefore play a role as a filler [27].

Concrete has many voids that are usually filled with water. The water can be classified depending on the type of void and degree of firmness with which the water is held in these voids. The water, which is responsible for the electrical conductivity of concrete, is called free water and is present in voids greater than 50 nm in size.

The transition zone, the region that exists between the hydrated cement paste and the aggregate, is very thin, with a thickness on the order of 10 to 50 μm . The transition zone is important in that it is the weakest zone in the concrete and thus influences its stiffness and durability.

The concrete used in the experiments is close to the one found at the Federal University of Minas Gerais. In addition, the dimensions of the four slabs were specified to meet the characteristics found in [28, 29]. In this way the mixture has the following characteristics: (1) cement Portland type I; (2) water/cement ratio 0.60; (3) cement/sand ratio 1 : 2.25. The inclusions have depth 35 mm from the surface of the concrete slab and have diameters of 19.05 mm as illustrated in Figure 4.

In order to get experimental results as close as possible to the numerical simulations it was prepared a GPR survey in four concrete slabs with the following characteristics: (1) without inclusion; (2) with metal inclusion; (3) with PVC inclusion; (4) water filled PVC inclusion. The survey was performed in a semi-anechoic chamber to avoid noise in the experiment according to the test setup illustrated in Figure 5.

The GPR equipment used in this work has its minimum configuration consisting of the following device grouped and used together:

- (a) 1.6 GHz or 2.3 GHz shielded antennas;
- (b) ProEx Control unity;
- (c) XV-11 Monitor with 1.2 m cable;
- (d) Cable X3M 4 m;
- (e) Battery Li-Ion 11.1 V/6.6 Ah.

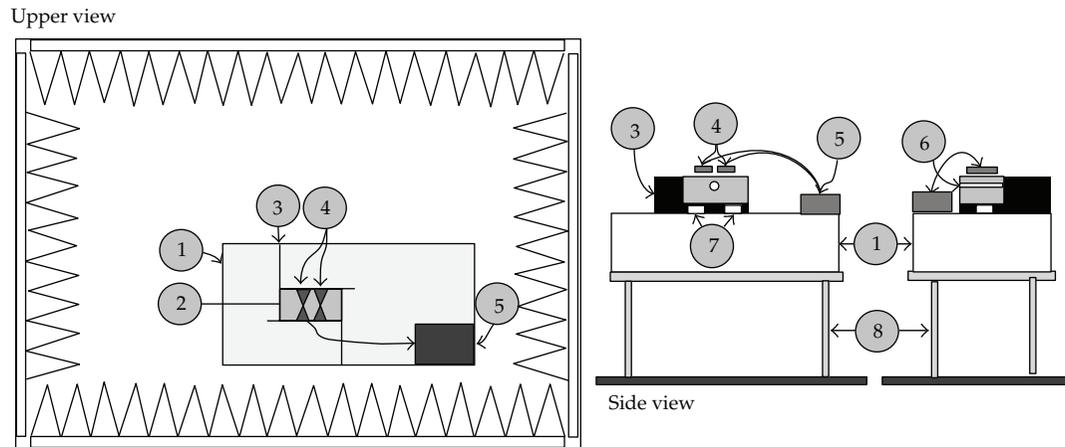


Figure 5: Schematic and photo of the assembly held in semi-anechoic chamber. (1) Polystyrene layer, (2) concrete block, (3) absorber plate, (4) bowtie antennas, (5) GPR control unit, (6) inclusion, (7) polystyrene blocks, and (8) wood table.

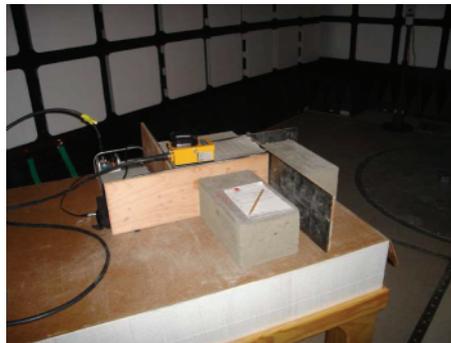


Figure 6: Experiment performed in a semi-anechoic chamber.

Prior to the GPR assessment of concrete in the chamber, test setup ambient levels (i.e., all equipment energized except the GPR equipment) were performed to verify 6 dB or more below the level required. The ambient measurements were performed using vertical polarization of a log periodic antenna.

The GPR equipment and its cable harness along with the concrete slabs lie on an insulated support 1000 mm (± 50 mm) above the floor of the test chamber. The dielectric constant of the insulated support was less than 1.4. The ground plane was not used for this experiment. The battery was located under the test bench. The equipment was at least 1000 mm from the chamber walls. No part of the GPR equipment was closer than 250 mm to the floor.

The GPR operated under typical loading and other conditions as in a regular survey to detect cracks and inclusions. The cables provided by the GPR equipment to perform the tests were shielded. Some cables that had excess lengths and were not bulk were bundled at the approximate center of the cable with the bundles 300 mm to 400 mm in length.

Figure 6 illustrates the experiment in the semi-anechoic chamber.

The climatic test conditions are defined in Table 5.

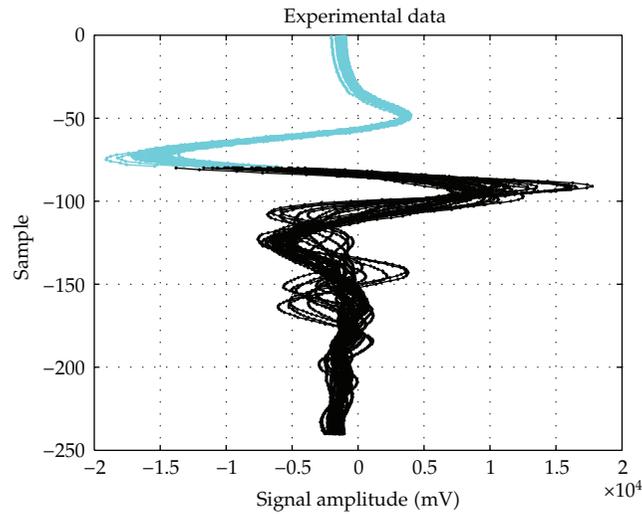


Figure 7: Experimental data of 1.6 GHz antenna: discarded samples in cyan and considered samples in black.

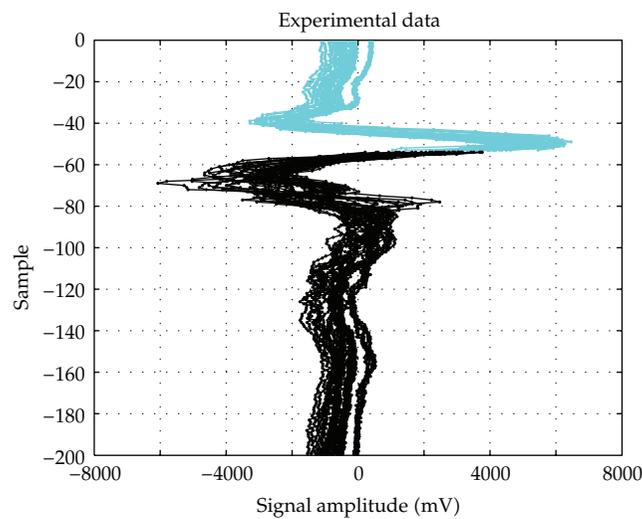


Figure 8: Experimental data of 2.3 GHz antenna: discarded samples in cyan and considered samples in black.

Based on simulated results and for higher number of samples, each concrete block was probed in its up and down sides. Therefore, 324 samples were collected with each antenna, 108 samples for each block inclusion (air, conductor, or PVC + water). As 12 samples of each block was taken over the inclusion, only these samples were considered to depth regression with half of them in top side and the other half in the underside of the blocks. The radius estimation was not experimented with real data set since all the inclusions in the concrete blocks had the same diameter (The experimental data set used can be downloaded in <http://www.enacom.com.br/doc/gprenacomsenai.zip>).

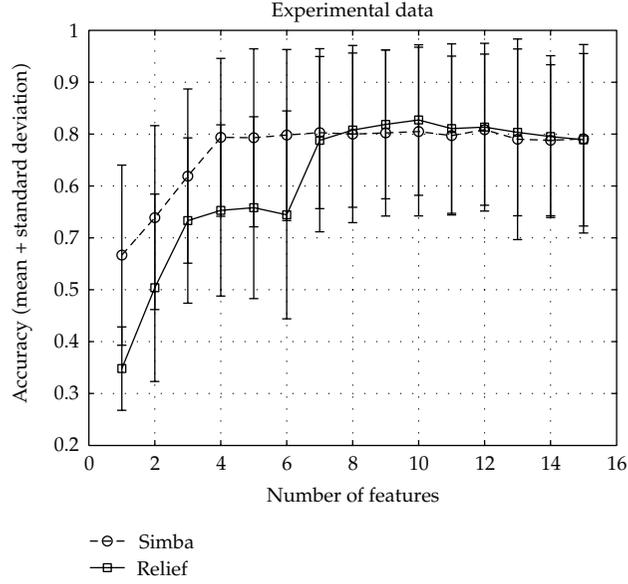


Figure 9: Accuracy results for k -NN with features extracted from experimental data of 1.6 GHz antenna to classify types of inclusions.

Each A-scan from both antennas had first echo excluded from data for models experiments. For 1.6 GHz antenna, just samples from 80th to 240th position were considered, as showed in Figure 7. And for 2.3 GHz, samples were selected from 54th to 200th position, as registered in Figure 8.

Before submitting these samples to classification and regression models, each data set was submitted to feature extraction as in [26] and after to Simba and Relief to rank features by their capacity of classes discrimination. The ranking results (from most to least important) for each antenna are as follows.

(i) 1.6 GHz antenna

(a) Simba: $\text{mean}(dW_s)$, fmFFT , $\text{std}(dW_s)$, mFFT , B_3 , SignalEnergy , B_2 , B_1 , $\text{std}(W_s)$, B_4 , $\text{max}(W_s)$, B_6 , B_5 , $\text{mean}(W_s)$, Delay .

(b) Relief: fmFFT , mFFT , $\text{std}(dW_s)$, SignalEnergy , $\text{std}(W_s)$, B_1 , $\text{mean}(dW_s)$, B_3 , B_6 , B_4 , B_2 , $\text{max}(W_s)$, Delay , B_5 , $\text{mean}(W_s)$.

(ii) 2.3 GHz antenna

(a) Simba: fmFFT , $\text{max}(W_s)$, $\text{mean}(dW_s)$, $\text{std}(dW_s)$, mFFT , B_6 , SignalEnergy , B_2 , $\text{std}(W_s)$, B_3 , B_5 , B_1 , B_4 , $\text{mean}(W_s)$, Delay .

(b) Relief: fmFFT , $\text{max}(W_s)$, $\text{std}(dW_s)$, B_3 , B_5 , B_2 , mFFT , SignalEnergy , $\text{mean}(dW_s)$, $\text{std}(W_s)$, Delay , B_6 , B_1 , $\text{mean}(W_s)$, B_4 .

Each of these features ranking was important to generate results to be compared to simulated data sets. As in the latter, no pre-processing of data was done before sample submission to models.

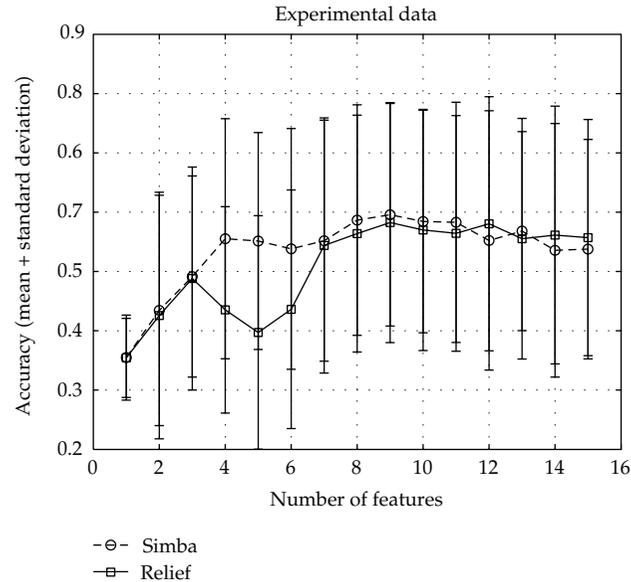


Figure 10: Accuracy results for k -NN with features extracted from experimental data of 2.3 GHz antenna to classify types of inclusions.

The $k = 3$ -NN for inclusion classification model used Squared Euclidean distance. And the accuracy was computed adopting $n = 50$ -fold cross-validation method adopting *cpartition* of MATLAB with its default values. The correspondent results are registered in Figure 9 for 1.6 GHz antenna and in Figure 10 for 2.3 GHz antenna.

Higher accuracies with fewer features were attained using the sequences obtained by Simba. But in both data sets the accuracies were smaller than expected. The results seen to be affected by real data problems like noise, clutter, and equipment instabilities as signal shifting. Then, more care should be taken in pre-processing steps addition to improve the quality of signals submitted to modeling step.

The PLP model was used with Log-sigmoid activation function as the regression model with two neurons used in internal layer. The percent error was computed using $n = 25$ -fold cross-validation method taking for each model with Hybrid Levenberg Marquardt allowing 10 epoch and 0.001 as maximum mean squared error (MSE). The expected results were rescaled to the allowed output range (between 0 and 1). The results for both antennas are shown, respectively, in Figures 11 and 12.

The features values extracted from both data sets permitted stable results despite the apparent higher standard deviation in the second figure because of the magnification done to visualization improvement. The features addition to this model indicated the existence of a small group of 6 features which should be considered to achieve smaller percent errors.

On the other hand, in both cases, the *Delay* feature, which obtained constant values for all samples, had significant impact in the chosen model, causing bad conditioned matrix of weights. The first models without considering this feature had significant percent error for depth estimation. But the subsequent models with *Delay* and other new features could not be trained satisfactory. So, these features should be disregarded until a new function for first delay arrival can effectively extract the corresponding information for samples discrimination.

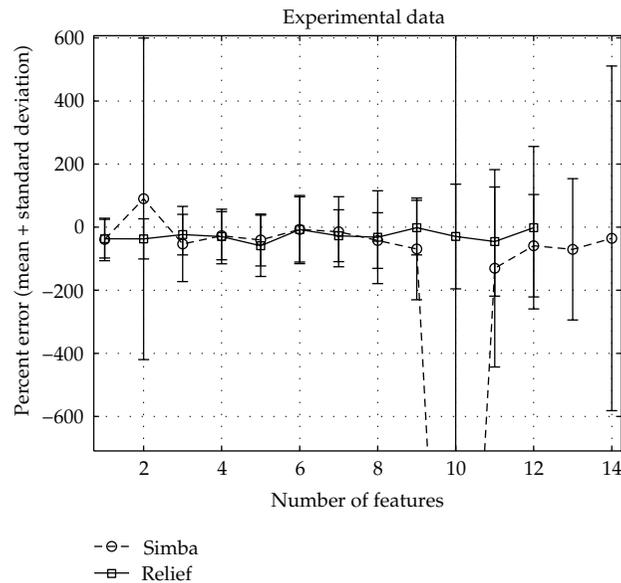


Figure 11: Percent Error main results for PLP with features extracted from experimental data of 1.6 GHz antenna to approximate depth of inclusions.

To complete the set of earlier results to be compared with the ones with simulated data sets, the same experimental data set and its extracted features were submitted to PCA and PLP-MGM of [30–32] for regression modeling. Using $n = 10$ -fold cross-validation method, PLP-MGM model with 3 internal neurons and saddle activation function was trained using Hybrid Levenberg Marquardt up to 10 epochs or until maximum MSE reaches below 0.001 level. For samples from A-scans, only components which contribute with more than 0.005% were considered in absolute percent error evaluation in models incrementing the number of components.

Before analyzing the results, it should be emphasized that the number of available samples for training and testing are much smaller than the ones simulated. Therefore, higher standard deviations are expected in the results considering the same number of folds in cross-validation method.

For 1.6 GHz antenna, 28 components were selected whereas for 2.3 GHz antenna, 35 components passed the permitted threshold. The respective results are exposed in Figures 13 and 14. The Pareto plots in each of these figures showed fewer significant components to explain the variance of the whole set. For both sets, the absolute percent mean error indicates small groups of no more than 10 principal components to achieve better estimation models for depth.

Using the extracted features, which were only 15, all of them were submitted to PCA and submitted to PLP-MGM models taking the same allowed parameters. These final results, which are in Figures 15 and 16, showed better estimation considering the same number of features was almost the same used for principal components, respectively. Based on these graphics, there really should exist a small set of features similar to the ones extracted which may obtain less complex models with satisfactory estimation even if nonrobust characteristics were considered.

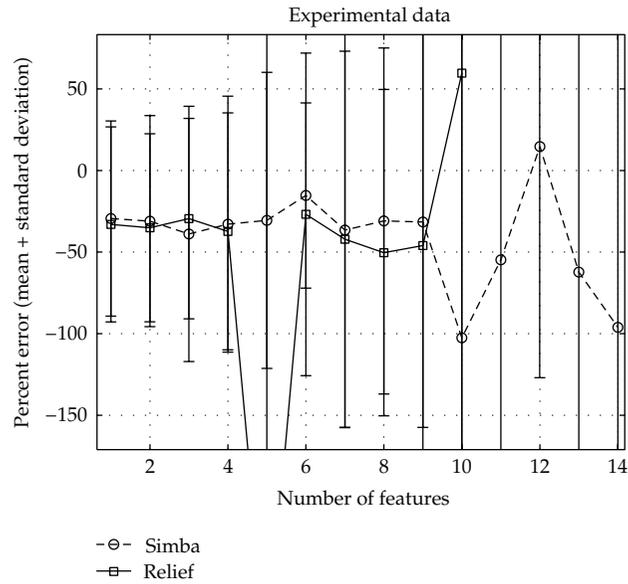
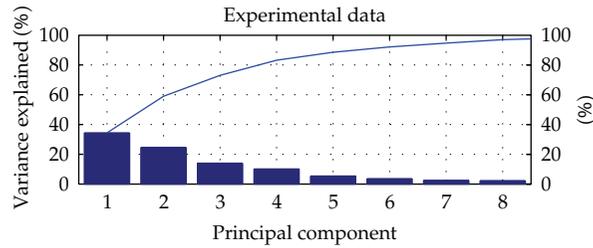
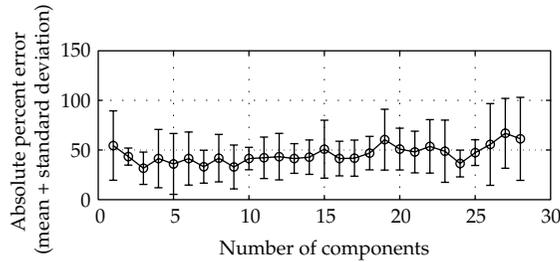


Figure 12: Percent Error main results for PLP with features extracted from experimental data of 2.3GHz antenna to approximate depth of inclusions.



(a)



(b)

Figure 13: Absolute Percent Error main results for PLP-MGM with PCA components of samples from experimental data of 1.6 GHz antenna to approximate depth of inclusions.

Table 3: Results of the relative error considering the depth prediction.

| Configuration | Max(L) | Train(s) | Test(s) |
|---------------|--------|----------|---------|
| PLP (286) | 0.04% | 8.44 | 0.016 |
| PLP (139) | 0.15% | 4.48 | 0.01 |
| PLP (51) | 20% | 6.2 | 0.011 |

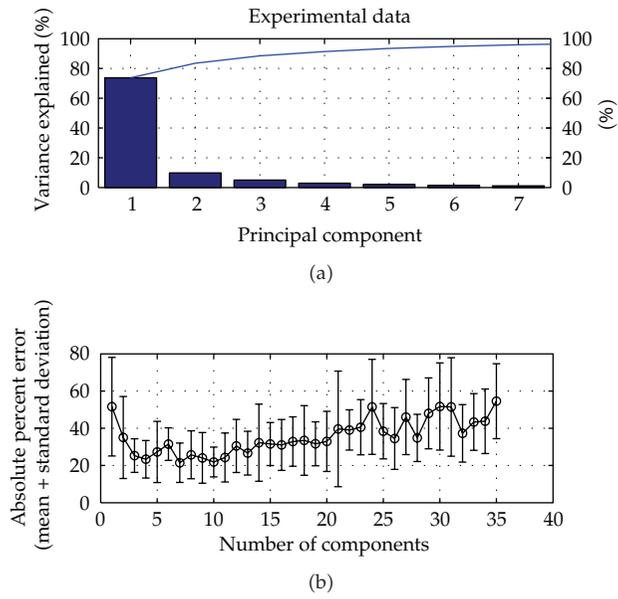


Figure 14: Absolute Error main results for PLP-MGM with PCA components of samples from experimental data of 2.3 GHz antenna to approximate depth of inclusions.

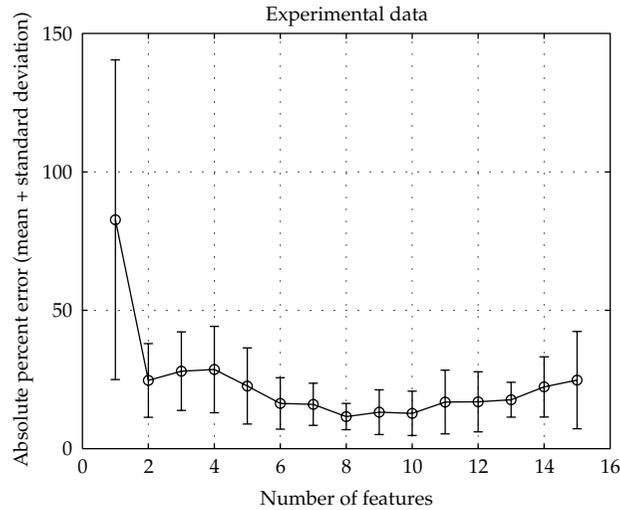


Figure 15: Absolute Percent Error main results for PLP-MGM with PCA components of features extracted from experimental data of 1.6 GHz antenna to approximate depth of inclusions.

However, models with extracted and selected features are preferable. Besides their better results, these latter models input components are more reliable since they consist of functions which behavior can be explained by the GPR operation and the related physics.

5. Conclusion

A number of important issues have been discussed in this paper. First the definition of estimation schemes for microwave imaging was outlined. The concept of inverse problem

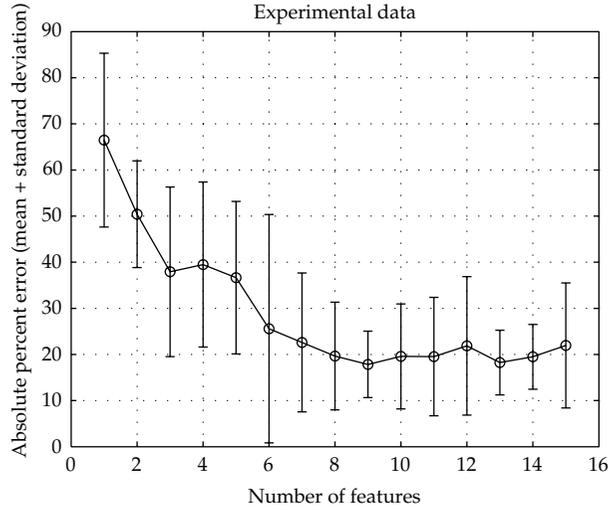


Figure 16: Absolute Error main results for PLP-MGM with PCA components of features extracted from experimental data of 2.3GHz antenna to approximate depth of inclusions.

Table 4: Results of the Relative Error considering the depth prediction given 286 dimensions, varying the number of neurons.

| Neurons | Mean(L) | Max(L) |
|---------|---------|--------|
| 6 | 0.009% | 0.27% |
| 7 | 0.008% | 0.16% |
| 8 | 0.003% | 0.05% |
| 9 | 0.001% | 0.04% |
| 10 | 0.004% | 0.09% |
| 11 | 0.004% | 0.09% |
| 12 | 0.005% | 0.10% |
| 13 | 0.006% | 0.15% |
| 14 | 0.005% | 0.16% |
| 15 | 0.008% | 0.10% |

for the GPR assessment of concrete was defined with respect to the inclusion reconstruction. Three 2D reconstruction algorithms were implemented and new configurations were proposed using Feature Selection. Numerical simulations with different types of inclusions were presented to assess the accuracy and efficiency of these algorithms for image reconstruction in real-world settings. Experiments were performed in a controlled environment to validate the results. Both data sets were submitted to PCA, feature extraction, and feature selection methods not only to improve the results, but also to obtain efficient explainable models.

As could be seen, a few extracted features are not robust to noise contamination and need some pre-processing to be reasonably evaluated. Some models were accordingly developed to reduce signal contamination, reduce the data formatting steps, or in some way overcome the large number of features and samples from A-scans. But the feature extraction process still has to be improved to avoid constant values, to gain a small general set of characteristics, and to reduce the pre-processing steps needed. In other words, it has to be developed a set of features to efficiently capture the information in which is expected to find

Table 5: Environmental test conditions.

| | |
|-------------|-----------------------------|
| Temperature | 23 ± 5 degrees C |
| Humidity | 20 to 80% relative humidity |

the elements capable of discriminate materials and explaining the phenomena involved in this inverse problem.

For most of the cases, the targets were satisfactorily reconstructed with the 2D algorithms and the convergence was assessed in terms of the relative error (see Tables 3 and 4). In conclusion, significant algorithmic flexibilities utilizing 2D based algorithm were demonstrated in terms of accommodating various forms of inclusions in single and iterative reconstruction. More informative and less complex models were obtained by the use of specific models to each inverse problem and trying to include some A-scans features that visually distinguish the materials and indicate inclusions.

Acknowledgment

This work was supported by Coordination for the Improvement of Higher Level—or Education—Personnel (CAPES), National Council for Scientific and Technological Development (CNPq), and Foundation for Research Support in Minas Gerais State (FAPEMIG), Brazil.

References

- [1] D. J. Daniels, *Ground Penetrating Radar*, The institute of Electrical Engineers, London, UK, 2004.
- [2] H. M. Jol, *Ground Penetrating Radar: Theory and Applications*, Elsevier Science, Oxford, UK, 2009.
- [3] L. B. Conyers, "Ground-penetrating radar techniques to discover and map historic graves," *Historical Archaeology*, vol. 40, no. 3, pp. 64–73, 2006.
- [4] M. M. Poulton, *Computational Neural Net-Works for Geophysical Data Processing*, Pergamon, Mysia, Turkey, 2001.
- [5] A. G. Davis, M. K. Lim, and C. G. Petersen, "Rapid and economical evaluation of concrete tunnel linings with impulse response and impulse radar non-destructive methods," *NDT and E International*, vol. 38, no. 3, pp. 181–186, 2005.
- [6] C. Maierhofer, H. W. Reinhardt, and G. Dobmann, Eds., *Non-Destructive Evaluation of Reinforced Concrete Structures: Non Destructive Testing Methods*, vol. 2, CRC Press, Boca Raton, Fla, USA, 2010.
- [7] K. J. Langenberg, K. Mayer, and R. Marklein, "Nondestructive testing of concrete with electromagnetic and elastic waves: modeling and imaging," *Cement and Concrete Composites*, vol. 28, no. 4, pp. 370–383, 2006.
- [8] S. R. Doctor, T. E. Hall, and L. D. Reid, "SAFT—the evolution of a signal processing technology for ultrasonic testing," *NDT International*, vol. 19, no. 3, pp. 163–167, 1986.
- [9] X. L. Travassos, D. A. G. Vieira, N. Ida, C. Vollaie, and A. Nicolas, "Inverse algorithms for the GPR assessment of concrete structures," *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 994–997, 2008.
- [10] X. L. Travassos, D. A. G. Vieira, N. Ida, and A. Nicolas, "In the use of parametric and nonparametric algorithms for the nondestructive evaluation of concrete structures," *Research in Nondestructive Evaluation*, vol. 20, no. 2, pp. 71–93, 2009.
- [11] J. H. Bungey, "Sub-surface radar testing of concrete: a review," *Construction and Building Materials*, vol. 18, no. 1, pp. 1–8, 2004.
- [12] X.-Q. He, Z. Q. Zhu, Q. Y. Liu, and G. Y. Lu, "Review of GPR rebar detection," in *Proceedings of the Progress in Electromagnetics Research Symposium (PIERS '09)*, pp. 804–813, Beijing, China, March 2009.
- [13] T. C. K. Molyneaux, S. G. Millard, J. H. Bungey, and J. Q. Zhou, "Radar assessment of structural concrete using neural networks," *NDT and E International*, vol. 28, no. 5, pp. 281–288, 1995.

- [14] M. R. Shaw, S. G. Millard, T. C. K. Molyneaux, M. J. Taylor, and J. H. Bungey, "Location of steel reinforcement in concrete using ground penetrating radar and neural networks," *NDT and E International*, vol. 38, no. 3, pp. 203–212, 2005.
- [15] D. Ayala-Cabrera, M. Herrera, J. Izquierdo, and R. Pérez-García, "Location of buried plastic pipes using multi-agent support based on GPR images," *Journal of Applied Geophysics*, vol. 75, no. 4, pp. 679–686, 2011.
- [16] M. Heiler, S. McNeil, and J. Garrett, "Ground-penetrating radar for highway and bridge deck condition assessment and inventory," in *Nondestructive Evaluation of Aging Bridges and Highways*, pp. 195–206, June 1995.
- [17] I. Abdel-Qader, F. Abu-Amara, and O. Abu-dayyeh, "Fractals and independent component analysis for defect detection in bridge decks," *Advances in Civil Engineering*, vol. 2011, Article ID 506464, 14 pages, 2011.
- [18] J. Wang, X. Xie, and H. Huang, "A fuzzy comprehensive evaluation system of mountain tunnel lining based on the fast nondestructive inspection," in *Proceedings of the International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE '11)*, pp. 2832–2834, 2011.
- [19] Z. M. Sbartai, S. Laurens, K. Viriyametantont, J. P. Balayssac, and G. Arliguie, "Non-destructive evaluation of concrete physical condition using radar and artificial neural networks," *Construction and Building Materials*, vol. 23, no. 2, pp. 837–845, 2009.
- [20] S. Bazarara, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley-Interscience, New York, NY, USA, 2006.
- [21] K. Kira and L. Rendell, "A practical approach to feature selection," in *Proceedings of the 9th International Workshop on Machine Learning*, pp. 249–256, 1992.
- [22] S. R. H. Hoole, "Artificial neural networks in the solution of inverse electromagnetic field problems," *IEEE Transactions on Magnetics*, vol. 29, no. 2, pp. 1931–1934, 1993.
- [23] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, *Feature Extraction-Foundations and Applications, Studies in Fuzziness and Soft Computing*, vol. 207, Springer, New York, NY, USA, 2006.
- [24] R. Gilad-Bachrach, A. Navot, and N. Tishby, "Margin based feature selection-theory and algorithms," in *Proceedings of the 21st International Conference on Machine Learning (ICML '04)*, pp. 337–344, July 2004.
- [25] S. Caorsi and G. Cevini, "An electromagnetic approach based on neural networks for the GPR investigation of buried cylinders," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 1, pp. 3–7, 2005.
- [26] F. A. A. Queiroz, D. A. G. Vieira, and X. L. Travassos, "Analyzing the relevant features of GPR scatteredwaves in time and frequency-domain," *Research in Nondestructive Evaluation*. In press.
- [27] P. K. Mehta, *Concrete: Structure, Properties and Materials*, Prentice Hall, Englewood Cliffs, NJ, USA, 1986.
- [28] O. Buyukozturk, T.-Y. Yu, and J. A. Ortega, "A methodology for determining complex permittivity of construction materials based on transmission-only coherent, wide-bandwidth free-space measurements," *Cement and Concrete Composites*, vol. 28, no. 4, pp. 349–359, 2006.
- [29] I. L. Al-Qadi, O. A. Hazim, and S. M. W. Suriad S.M., "Dielectric properties of Portland cement concrete at low radio frequencies," *Journal of Materials in Civil Engineering*, vol. 7, no. 3, pp. 192–198, 1995.
- [30] D. A. G. Vieira, L. Travassos, R. R. Saldanha, and V. Palade, "Signal denoising in engineering problems through the minimum gradient method," *Neurocomputing*, vol. 72, no. 10–12, pp. 2270–2275, 2009.
- [31] X. L. Travassos, D. A. G. Vieira, N. Ida, C. Vollaie, and A. Nicolas, "Characterization of inclusions in a nonhomogeneous GPR problem by artificial neural networks," *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 1630–1633, 2008.
- [32] X. L. Travassos, D. A. G. Vieira, V. Palade, and A. Nicolas, "Noise reduction in a non-homogenous ground penetrating radar problem by multiobjective neural networks," *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1454–1457, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

