In the use of parametric and non-parametric algorithms for the non-destructive evaluation of concrete structures

L. Travassos [a,b] , D. A. G. Vieira [a] , N. Ida [c] , and A. Nicolas [a]

[a] *AMPERE Lab, Ecole Centrale Lyon, France*

[b] *SENAI CIMATEC, Brazil*

[c] *University of Akron, USA*

Abstract

In order to maximize the performance of subsurface radar investigations there exists a vast range of algorithms that can be applied to the data obtained to find the target's location and reflectivity. Minimizing the data computing time becomes the primary issue in choosing one algorithm over another. In this work, three imaging methods are used to reconstruct the geometric and dielectric characteristics of buried cylinders. The subject of this work is qualitative comparisons between these methods based on different performance parameters with respect to the discrepancy between calculated and true object characteristics. These methods are applied to the detection of inclusions in concrete structures.

Key words: Numerical modeling, Inverse Problems, Non-destructive testing, Neural Networks, Particle Swarm Optimization.

1 Introduction

Image classification is a commonly pursued area in diverse fields such as military, security systems, health monitoring and biomedical engineering. This is due to the necessity of eliminating the risk of human misinterpretation by using a machine. The main idea is to obtain information by processing data obtained from sensors. Such machines can substantially reduce the time employed for interpretation and improve accuracy of decision making by human operators.

Email address: lucas.travassos@ec-lyon.fr (L. Travassos).
URL: http://www.ampere-lab.fr (L. Travassos).

The major technical drawback to apply imaging algorithms is the large variation in the inspected target signatures due to environmental conditions, geometric variations, noise, and sensors' characteristics. Therefore this can be considered as a multidisciplinary problem requiring contributions from diverse technologies.

Notwithstanding the importance of the above-mentioned applications, this paper concentrates in the Non-Desdructive Testing (NDT) of concrete structures, which requires reliable measurement methods. Pulsed radars are attractive as environment measurement methods for various applications including the examples above. The waveform data is obtained by scanning an omnidirectional antenna. The use of this waveform for estimating target characteristics is known as an ill-posed inverse problem.

In the past, various imaging or inversion techniques have been developed to refocus the scattered signals back to their true spatial location. Most of them were based on the numerical inversion of integral equations. All these techniques are characterized by a high level of complexity and, even in the cases which accuracy of the solution is good, they result in a significant computational burden. Consequently, 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.

Therefore, it has become necessary to use more efficient analysis for the raw-data interpretation. Such analysis requires algorithms by which problems having complex scattering properties can be solved as accurate and as fast as possible. This specification is difficult to achieve when dealing with iterative algorithms characterized by a forward solver as part of the loop, which often makes the solution process computationally prohibitive for large problems. An alternative approach is the use of model-free methods based on example data. This category is represented by Artificial Neural Networks (ANN).

Three imaging methods are investigated in this paper. First, a reverse-time migration algorithm is implemented to refocus the target to its true spatial location. Also, a model fitting technique was used to classify the target using Particle Swarm Optimization. Finally, buried inclusions characteristics considering a non-homogenous host medium were found by using ANN. They are examined with respect to complexity and the ability to reproduce the characteristic of the underground targets considering the host medium as a concrete structure.

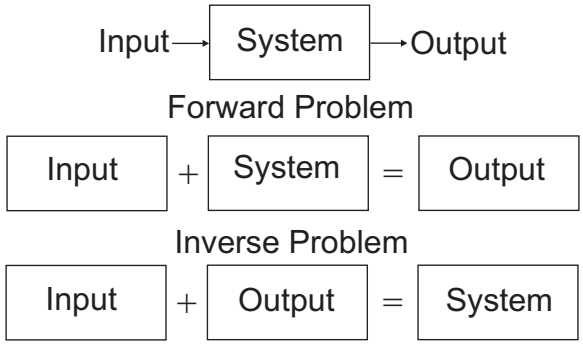

Fig. 1. Forward and Inverse problems.

2 Inverse problems

To predict the result of a measurement it is required a model of the system under investigation, and a physical theory linking the parameters of the model to the parameters being measured. This prediction of observations, given the values of the parameters defining the model, constitutes the forward problem depicted in Fig. 1. The inverse problem consists in using the results of actual observations to infer the values of the parameters characterizing the system under investigation.

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. A problem is called ill-posed if it satisfies one of the following three conditions.

(1) Existence. There may be no model that exactly fits the data. This can occur in practice because our model of the system's physics is approximate or because the data contain noise.
(2) Uniqueness. If exact solutions do exist, they may not be unique, even for an infinite number of exact data points. That is, there may be other solutions besides the one found that exactly satisfy the expected input.
(3) Instability. The process of computing an inverse solution can be extremely unstable in that a small change in measurement can lead to an enormous change in the estimated model. The condition of stability is often violated for ill-posed problems.

The focus here is on the inverse problem of finding the characteristics of a target buried in a dielectric material given the reflected field measured by the antenna. 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 radii) 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 or evolutionary algorithms. One major limitation of using parametric algorithms is the time of calculation that can be prohibitive for 3D problems.

Non-parametric 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. Migration algorithms belong to non-parametric algorithms. Migration is an imaging technique commonly used for seismic prospecting. However, it can be applied for general media and targets. Considering the radar inspection of concrete structures a migration algorithm for vector waves was implemented based on the idea of matched filter. This algorithm is described next.

3 Reverse-time migration algorithm (RTMA)

The migration process essentially constructs the target reflector surface from the recorded surface. The migration technique has been widely developed in acoustic, seismic and geophysical engineering and was originally developed in two-dimensional form by Hagedoorn (1954). Some surveys can be found in (12; 13; 15). In (1) a pair of FDTD reverse-time migration algorithms was presented for radar data processing that use linear inverse scattering theory to develop a matched-filter response for the radar problem. These algorithms were developed for both bistatic and monostatic antenna configurations.

Spatial location and reflectivity are typical information obtained from radar. Since most radars use broad beamwidth antennas, the energy reflected from a buried structure is recorded over a large lateral aperture in the image space. For example, a monostatic survey can be used to collect the data over a discrete object, such as a pipe in which the diffraction appears as a hyperbola in the space-time image. In this case, no further processing may be needed if the goal is simply to detect the pipe. However, imaging algorithms must be used to move the observed scattering events to their true spatial location and to

estimate the target's reflectivity.

The algorithm developed in (1) is based on the notion of a matched filter, which is widely used in radar applications. The matched filter concept can be explained as a correlation of the received signal with the expected or estimated signal from a specific target. If this correlation produces a large value, then it is likely that the target is present. The resulting algorithm can be directly related to reverse-time migration. Using this technique, an image can be perceived as a backpropagated wave-field reconstruction of the dielectric contrast within the ground.

Mathematically, the matched filter transfer function H is expressed as the complex conjugate of the expected received waveform due to the target which the filter is being matched. The output of the matched filter for N transmitters, located at the position vectors $\mathbf{r}_n^{''}$, and an M-element receiver array, located at the position vectors $\mathbf{r}_m^{''}$ is expressed as (1)

$$S\left(\mathbf{r}\right) = \sum_{n=1}^{N} \int \sum_{m=1}^{M} H\left(\mathbf{r}_n^{''}, \mathbf{r}^{'}, \mathbf{r}_m; \omega\right) U_n\left(\mathbf{r}_m; \omega\right) d\omega \tag{1}$$

where U_n is the received waveform due to the nth transmitter as shown in the Fig. 2.

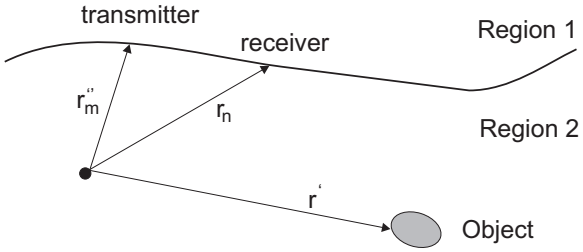

Fig. 2. Radar problem geometry.

The problem geometry depicted in Fig. 2 consists of two half spaces where Region 1 corresponds to free space. An inhomogeneous ground, characterized by constitutive parameters μ_0, $\epsilon_2(r)$, is denoted as Region 2. Considering a weakly scattering object of finite size with constitutive parameters μ_0, $\epsilon(r)$ located within the ground the matched filter will maximize the output power at $t = 0$ if the complex conjugate of the matched filter is equal to the received signal.

The scattered electric field is expressed as

$$\mathbf{E}_{sca}\left(\mathbf{r}\right) = -\int \overline{\mathbf{G}}\left(\mathbf{r}, \mathbf{r}^{'}\right)\left[k^2\left(\mathbf{r}^{'}\right) - k_b^2\left(\mathbf{r}^{'}\right)\right]\mathbf{E}\left(\mathbf{r}^{'}\right) d\mathbf{r} \tag{2}$$

where $\overline{\mathbf{G}}\left(\mathbf{r}, \mathbf{r}^{'}\right)$ is the background dyadic Green's function, and k and k_b are the

ordinary wavenumber and the Born-approximated wavenumber, respectively. If the scatterer is small enough, the following equation holds approximately

$$\mathbf{E}_{sca}\left(\mathbf{r}\right) = -\omega^2\mu_0\overline{\mathbf{G}}\left(\mathbf{r},\mathbf{r}^{'}\right)\mathbf{E}_{inc}\left(\mathbf{r}^{'}\right) \tag{3}$$

where E_{inc} indicates the incident field. By using this approximation, the matched filter H can be written as

$$H\left(\mathbf{r},\mathbf{r}^{'},\mathbf{r}^{''}\right) = \mathbf{u}_r^*\left\{-\omega^2\mu_0\overline{\mathbf{G}}\left(\mathbf{r},\mathbf{r}^{'}\right)j\omega\mu_0\overline{\mathbf{G}}\left(\mathbf{r}^{'},\mathbf{r}^{''}\right)T\mathbf{u}_t\right\} \tag{4}$$

where $*$ denotes the complex conjugate, and $\mathbf{u}_t$ is the receive antenna effective length. The final image is then found by applying the complex conjugate of the measured data to the filter

$$S\left(\mathbf{r}'\right) = \left\{j\omega\mu_0\overline{\mathbf{G}}\left(\mathbf{r}^{'},\mathbf{r}\right)\left[-j\omega R\mathbf{u}_r\right]^*\right\}\left\{j\omega\mu_0\overline{\mathbf{G}}\left(\mathbf{r}^{'},\mathbf{r}^{''}\right)T\mathbf{u}_t\right\}. \tag{5}$$

Equation (5) gives the migrated data as a function of frequency and the antenna locations. The first term is the electric field generated by a current source $\left[-j\omega R\mathbf{u}_r\right]^*$. If the time dependency of the received signal is introduced, this source is expressed as a derivative and time reversal $\left[-j\omega R\mathbf{u}_r\right]^* \Rightarrow R^{'}\left(-t\right)\mathbf{u}_r$. The field generated by this source will be refereed to as the backpropagated electric field $\mathbf{E}_{bp}$. The second term is simply the incident field $\mathbf{E}_{inc}$. Reintroducing the frequency dependencies and referring to Eq. (1), a complete expression for the migrated data is now shown as

$$S\left(\mathbf{r}'\right) = \sum_{n=1}^{N}\int\sum_{m=1}^{M}\mathbf{E}_{mn,bp}\left(\mathbf{r}^{'};\omega\right)\mathbf{E}_{n,inc}\left(\mathbf{r}^{'};\omega\right)d\omega \tag{6}$$

$$\mathbf{E}_{mn,bp}\left(\mathbf{r}^{'};\omega\right) = j\omega\mu_0\overline{\mathbf{G}}\left(\mathbf{r}_m,\mathbf{r}^{'};\omega\right)\left[-j\omega R_n\left(\omega\right)\mathbf{u}_r\right]^* \tag{7}$$

$$\mathbf{E}_{n,inc}\left(\mathbf{r}^{'};\omega\right) = j\omega\mu_0\overline{\mathbf{G}}\left(\mathbf{r}^{'},\mathbf{r}_n^{''};\omega\right)T\left(\omega\right)\mathbf{u}_t \tag{8}$$

where the subscripts m and n denote the field or signal due to the nth transmitter and mth receiver. These equations are now applied to bistatic survey depicted in Fig. 3 using the FDTD technique.

The dielectric medium in this problem is simulated with electrical characteristics of concrete (4) with a mean relative electrical permittivity value of 6 and standard deviation 0.25, i.e, a non-homogeneous medium. The FDTD forward solver is used with a source defined by a differentiated gaussian pulse with a center frequency of $f = 900MHz$ and bandwidth between 0.3 and $2GHz$ (Fig. 4) with antennas being simulated as dipoles of infinite extension in the z direction. A sampling interval of 20 ps was chosen to meet the stability criteria

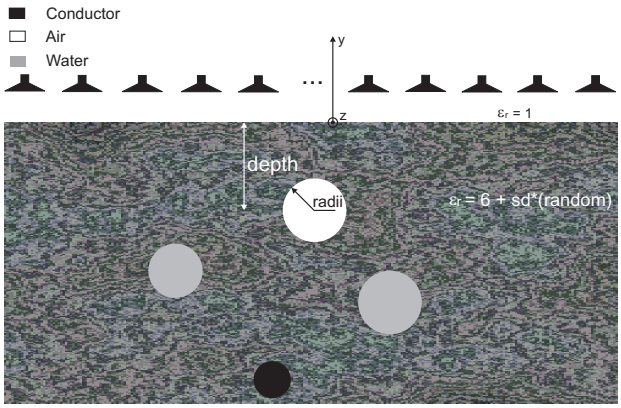

Fig. 3. Configuration consisting of a circular cylinders located in a non-homogeneous dielectric. The problem is composed by a set of inclusions with different properties. The aim is, having the scattered wave, define the inclusions geometry and properties.

of the FDTD, while 1500 samples were collected for each trace corresponding to a time interval of 30 ns. In order to control the numerical dispersion and provide a good discretization for the inclusions, the spatial steps were chosen as $\Delta x = \Delta y = 12mm$. The aim of this problem is, given an incident wave, and scattered wave, to determine the physical and spatial characteristics of the inclusion.

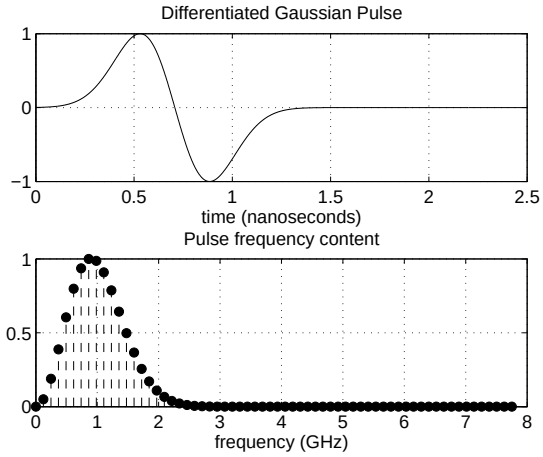

Fig. 4. Pulse description.

The implementation in FDTD is accomplished by propagating the incident field in reverse while simultaneously propagating the back propagated field forward. As an example, consider the FDTD scenario consisted of four cylinders (with different radius), conductor (1), air (2) and water (1), buried in concrete. The raw data, collected at 31 receiver locations separated by 7.2 cm with the transmitter located above the dielectric slab. The three traces near the transmitter were zeroed out since the receiving and transmitting antennas were in too close proximity. The final image occurs where the incident field intersects the backpropagated field.

The backpropagated field is the reference field simulated with a non-homogeneous medium. The matched filter used an homogeneous medium with $\epsilon = 6$ and $\sigma = 1mS$. Three simulations with the incident field in dielectric's far left, right and center were done.

Figure 5 shows the final image obtained at $t = 0$ that provides a more exact location of the inclusions positions. This process improves the final images provided by the radar inspection data. In addition its implementation is very simple.

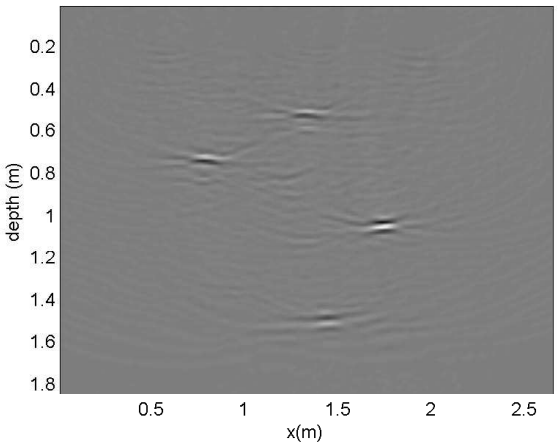

Fig. 5. This figure shows the bistatic reverse-time migration; final data. This is the filtered scattered wave, where the dots represent the center of the inclusions.

Migration algorithms belong to the class of non-parametric algorithms. 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. Even though this is a robust method for solving the above mentioned problem, it can only identify the number of inclusions and their centers. This algorithm assumes that the background medium is known - a common assumption in many commercial softwares for the interpretation of ground penetrating radar (GPR) data (5). In addition, in this simulation line sources were used which can create a problem when the objective is to find targets in close proximity. In real problems, this algorithm can be used to locate conductors and buried pipes but it can not find the exact characteristics of the inclusions.

Another possible technique to solve this problem is the use of parametric algorithms which are based on optimization algorithms. The parametric approach will be briefly described in the following section.

4 Model fitting

Model fitting method is one of the parametric imaging algorithms which applies 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. Thus, it can be written as an optimization problem involving the scattered wave of the unknown object $\mathbf{E}(\theta_0)$, the reference object, and the scattered wave of a test object $\mathbf{E}(\theta)$. Thus, θ^*, the optimal θ, is the argument that minimizes the error of the reference object scattered wave $\mathbf{E}(\theta_0)$ relative to the test object scattered wave $\mathbf{E}(\theta)$. Mathematically:

$$\theta^* = \ arg\ min\ f(\theta) = \sum_{i=1}^{ns} (\mathbf{E}(\theta_0) - \mathbf{E}(\theta))^2 \tag{9}$$

where ns are the sample points where the scattered wave is measured. Note that $E(\theta_0)$ is known even though θ_0 is unknown, it is measured at the receiver antennas.

The scattered $\mathbf{E}(\theta)$ is then generated assuming one test θ, and the optimization procedure aims at minimizing the error between $\mathbf{E}(\theta_0)$ and $\mathbf{E}(\theta)$ in such a way to identify the scatter object θ_0. This paper assumes that θ, and θ_0, are composed by a set of cylindrical inclusions, thus, $\theta = \{(r_1, c_1, e_1, \sigma_1), ..., (r_n, c_n, e_n, \sigma_n)\}$.

The parametric approach can robustly define the number of inclusions n and their centers c_i, but it cannot find the other variables. This work uses the parametric approach to define the radii and the physical properties of the inclusions and, then find the other properties using the parametric approach.

The problem described in (9) is usually multi-modal, as shown in Fig. 6, where the unknowns are the radius 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.

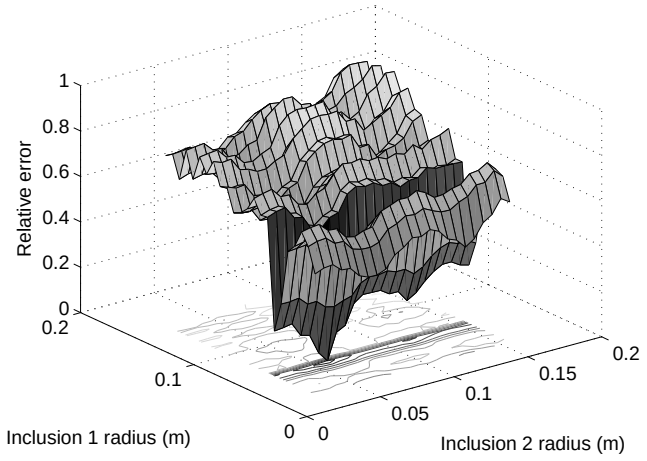

Fig. 6. Surface plot of Eq. (9) representing the problem of finding the radius of two inclusions, given that their location and physical properties are known. Even in this simple example the surface is clearly multi-modal, thus, stochastic methods must be employed in the parametric approach.

4.1 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is one of the latest evolutionary optimization methods inspired by nature (2)(3) that includes evolutionary strategy (ES), evolutionary programming (EP), genetic algorithm (GA), and genetic programming (GP). PSO is based on the metaphor of social interaction and communication such as bird flocking and fish schooling. PSO is distinctly different from other evolutionary-type methods in that it does not use the filtering operation (such as crossover and/or mutation), and the members of the entire population are maintained through the search procedure so that information is socially shared among individuals to direct the search towards the best position in the search space.

The original intention was to graphically simulate the choreography of a group of birds. At some point in the development of the model it was noticed that the algorithm was in fact an optimizer. Using an empirical process many parameters were eliminated - the ones relevant to the social model - but not to the optimization, resulting in a simple algorithm.

In a PSO algorithm, each member is called a particle, and each particle moves around in the multi-dimensional search space with a specific velocity. According to the global neighborhood, each particle moves towards its best previous position and towards the best particle in the whole swarm.

The PSO 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" in the search space. For each particle of the swarm during the iterations, the positions of the best solution found to a given particle, called pbest (particle best) is saved. The best value found considering all the particles 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. The PSO algorithm can be described as following:

(1) Initialize the swarm of particles with random positions and velocities.
(2) For each particle calculate the objective function.
(3) Find *pbest*.
(4) Find *gbest*.
(5) Change the velocity and position of each particle according to Eqs. 10 and 11

$$v = v + c_1 * rand * (pbest - x) + c_2 * rand * (gbest - x) \tag{10}$$

$$x = x + v \tag{11}$$

(6) Return to step 2 untill a stop criterion is achieved.

The velocity of each particle in each dimension is limited by a maximum velocity, $Vmax$. The acceleration constants c_1 and c_2, represent the trade-off between the search in the direction of *pbest* and *gbest*. Usual values for c_1 and c_2 are equal to 2 and $Vmax$ between 10% and 20% of the variable range in each dimension.

The proposed hybrid algorithm was used to find the inclusions' characteristics given a non-homogenous medium with standard deviation $sd = 0.25$. The inclusions are defined in Table 1. The search space considered $radii = [5\ 10]cm$ and each inclusion could be one of the following materials: air, conductor and water.

Once the inclusions' positions were found using the migration algorithm this information was used as one additional parameter for the model fitting technique. The hybrid approach that combines migration with PSO is compared with the PSO alone for problems with one, two and three inclusions. Table 2 presents the results in terms of the number of function evaluations, taken over an average of 10 simulations, to define the inclusion. Each function evaluation requires about $12s$ of computation.

The results show that the number of function evaluations decreases significantly when the approach proposed in this paper is applied compared to the traditional parametric approach. It is important to remark that the migration algorithm could not solve the whole problem alone.

Table 1
Inverse scattering inclusion definitions

	Center(x,y) [m]	Radii [cm]	Material
Inc_1	(0.78,0.84)	6.6	Air
Inc_2	(1.32,0.60)	9.0	Conductor
Inc_3	(1.74,0.60)	9.6	Water

Table 2
Average number of FDTD evaluations, considering 10 tests, by the hybrid approach proposed in this paper compared with the PSO alone.

Num. of Inclusions	Hybrid	PSO
1	41	670
2	475	3800
3	2100	NA

For instance, in a problem with 2 inclusions, the PSO has to find 4 variables, while the hybrid approach proposed requires only 2. This makes the search more efficient and, therefore, faster. Another remarkable fact is that the cost associated increases quickly as the number of inclusions increases.

This fact limits the use of our approach in situations when the number of inclusions is high. For the pure PSO approach the number of function evaluations diverges with only 3 inclusions, and it is marked as Not-Attainable, NA, in Table 2. However, the large search region outlined in this problem makes solving the inverse problem difficult. Although iterative optimization approaches require long computing time, these approaches provide much better image qualities for high-contrast objects than linear inversions such as diffraction tomography (6)(7).

The problems presented here show favorable results but real-world applications require "black-box" solutions. This can be done using Artificial Neural Networks (ANN) to improve the detection of buried cylinders. The next section describes the use of the Parallel Layer Perceptron (PLP) (10) architecture for this purpose. A new approach using PLP and Principal Component Analysis (PCA) is proposed to compress the radar data in order to train ANN faster.

5 Neural networks

The brain can be viewed as a highly complex, nonlinear parallel computer that can compute in an entirely different way compared with a conventional

computer. The brain is a complex network with approximately 10^{10} neurons having over $6 \cdot 10^{13}$ interconnections. Neural events occur at millisecond speeds whereas events in common computers occur in less than nanoseconds. However, the brain can compensates this slow speed through its massive number of neurons and interconnections. For example, the human brain can recognize a familiar face embedded in an unfamiliar scene in 100-200 ms, whereas a computer can take some hours.

Recognition of the brain's impressive power has led to interest in the development of Artificial Neural Network (ANN) technology. ANN are parallel computational models comprised of densely interconnected adaptive processing units. A very important feature of this technology is its adaptive nature, whereby the problem is solved by feeding the system with examples as in the human brain. This feature makes such computational models very appealing in application where one has little or incomplete understanding of the problem to be solved but where training data are readily available. Another key feature is the intrinsic parallel architecture that allows fast computation of solutions.

ANN can be used in a wide range of applications including pattern classification, speech synthesis and recognition, adaptive interfaces between humans and complex physical systems, function approximation, image compression, associative memory, clustering, forecasting and prediction, combinatorial optimization and nonlinear system modeling, and control. In the context of this work the ANN are used in Intelligent Signal Processing (ISP) which is characterized by the use of model free (intelligent) methods based on training data. In addition, ISP implies the ability to extract system information from the example data alone and is less dependent on a priori environmental and system information or simply, is less unstable.

A formal definition of an ANN according to (16) is: "A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

(1) Knowledge is acquired by the network through a learning process.
(2) Interneuron connection strengths known as synaptic weights are used to store knowledge."

Neurons are the basic building blocks that make up the ANN. They are usually made to be all similar and can be interconnected in various ways to make a network. The ANN achieves its ability to learn and then recall that learning through the weighted interconnections of those neurons. The interconnection architecture can be very different for different types of networks. Architectures can vary from different types of feedforward to recurrent structures.

The study focused on the development and performance of a comprehensive

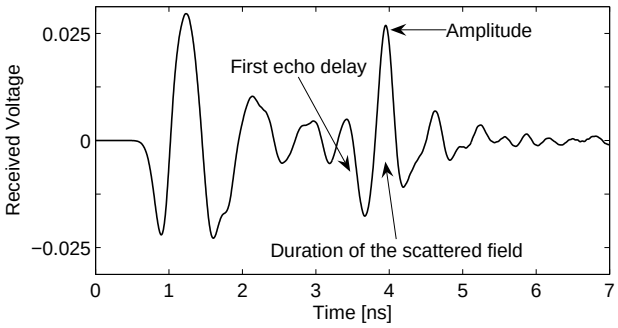

Fig. 7. Reflected wavefield from a buried target.

ANN model for the analysis of jointed concrete slabs are presented in (18) and (19). The inverse scattering problem using parallel networks and networks with multiple outputs for an homogenous host medium was presented in (8) and (9). In (9) it is shown that both configurations could deliver reasonable and very similar results using input parameters from the scattered wave defined as: 1) the peak amplitude of the reflected field; 2) the delay of the first reflected echo, calculated with respect to the time of arrival, at the receiving point, of the direct field); 3) a measure of the duration of the scattered field (see Fig. 7). This paper considers the case of non-homogenous host medium and, surprisingly, using a network with multiple outputs and parallel networks were not adequate to estimate the inclusion geometry.

Therefore, to solve a non-homogenous problem the parameters presented in (9) are not sufficient. To overcome this difficulty, we implemented an algorithm to "squeeze" the scattered wave in order to collect more information about the scatterer itself using Principal Component Analysis (PCA).

The scattered wave is obtained using FDTD. The problem can be summarized as the use of an incident wave and a scattered wave 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 paper addresses a 2D problem where a cylinder of unknown characteristics is buried in a non-homogeneous dielectric. The incident and scattered wave are simulated using FDTD to train the ANN. The dielectric medium uses the electrical characteristics of concrete (4) with a mean relative electrical permittivity value of 6 and standard deviation 0.15, i.e, a non-homogenous medium. The investigation domain is illuminated by a differentiated gaussian pulse depicted in Fig. 4. In order to control the numerical dispersion and provide good discretization for the inclusions the spatial steps were chosen as $\Delta x = \Delta y = 6mm$. The aim of this problem is, given an incident wave, W_i, and scattered wave, W_s, to determine the *radii* and *depth* of the inclusion.

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

The ANN has been trained with a set of different inclusion examples, constructed by varying the $radii$ in the range $[0.02 \div 0.1]$ m according to the rule $radii = 0.02 + i \times 0.001$, $i = 0, .., 80$, with ϵ_r in the range $[1 \div 10]$, according to the rule $\epsilon_r = 1 + i \times 1$, $i = 0, ..., 9$, σ in the range $[0 \div 4000]$ S/m according to the rule $\sigma = 0 + i \times 500$, $i = 0, ..., 8$ and $depth$ in the range $[0.05 \div 0.25]$ m according to the rule $depth = 0.05 + i \times 0.025$, $i = 0, ..., 9$ 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, i.e., the learning process becomes slower and less effective. In the next section the curse of dimensionality will be introduced and, afterwards, it will be followed by the Principal Component Analysis (PCA) as an effective way to reduce the dimensionality.

5.1 The curse of dimensionality

One phenomenon that takes place in high dimensional data is the sparsity of the sample points (20). Given a data set S with T data points uniformly distributed in a p-dimensional unit sphere centered at the origin, the median distance given from the origin to the closest sample is then given by:

$$d(p, T) = \left(1 - \frac{1^{1/T}}{2}\right)^{1/p}. \tag{12}$$

In a sample size of 1640 and 1200 dimensions, $d(1200, 1640) = 0.9935$, which means that the samples are closer to the boundary of the space than to any other data point. Moreover, for an ANFIS topology the number of rules R for a system with p inputs and P premisses is $R = P^p$, hence it increases exponentially with the dimension p, which makes the learning slow (14). The following section presents the Principal Component Analysis, which will be used for dimensional reduction.

5.2 Principal Component Analysis

In some situations, the dimension of the input vector is large, but the components of the vectors are highly correlated (redundant). Principal Component

Analysis (PCA) is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. The main advantage of PCA is that once these patterns in the data are found, the data can be compressed, *i.e.* by reducing the number of dimensions, without much loss of information. This technique is commonly used in image compression.

This technique has three effects: it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other), it orders the resulting orthogonal components (principal components) so that those with the largest variation come first, and it eliminates those components that contribute the least to the variation in the data set.

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 $S_q = \{x_1, ..., x_T | x_i \in \mathbb{R}^p\}$ to construct a new data set $S_p = \{y_1, ..., y_T | y_i \in \mathbb{R}^q\}$, where $q \leq p$.

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 considered here is initially in $\mathbb{R}^{1200}$. It can be projected in a $\mathbb{R}^{286}$ without any loss of information, *i.e.*, 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.

5.3 The expected error

The model structure problem is given by choosing among a set of possible functions $f(w, x)$, $w \in \Lambda$ [1], the possible Parallel Layer Perceptron (PLP) in the present case, the one that optimizes a given quality criterion. Mathematically

$$w^* = \arg \min_w R(f(w, x)) \tag{13}$$

[1] In the PLP case Λ is the set of Real numbers but, in general, it can be a set of scalar quantities, a set of vectors, or even a set of abstract elements (22).

where $R(.)$ [2] is a pre-defined quality criterion and w^* is the argument that minimizes Eq. 13. To apply the selected model structure, it is required to define the best approximation of the desired output, d, therefore, a measure of discrepancy, $L(.)$, between the desired and obtained outputs should be employed. The expected risk (error) between the desired and the approximate outputs can be expressed (22) as:

$$R(w) = \int L(d, f(x, w))dF(x, d). \tag{14}$$

The expected risk $R(w)$ measures the expected test error for the neural network, i.e., the ANN performance. The aim of the machine the learning procedure aims to find $f(w, x)$, $w \in \Lambda$, that minimizes the risk functional in Eq. 14. The integral cannot be evaluated directly since the distribution $F(x, d)$ is unknown and the only available information is the training set. The training set is made of T random vectors $z_i = (x_i, d_i)$, $i = 1, ..., T$, independently and identically distributed $(i.i.d)$ according to some unknown but fixed probability distribution $F(z)$. The training set S can be written as:

$$S = ((x_1, d_1), ..., (x_T, d_T)). \tag{15}$$

The risk $R(.)$ defined in Eq. (14) can be asymptotically approximated, given some consistence conditions (22), when the number of training set samples T tends to infinity. Of course such an infinite size set is not available. To overcome this problem, resampling techniques can be used to approximate the expected risk.

The simplest resampling technique is the Hold Out (HO), also called external validation or simply validation. It consists of removing samples from the initial learning set, and using them for validation. That was used in many works, by dividing the set in training and test set. In this paper we employed the k-fold cross-validation. For k-fold cross-validation, the training data S is divided into k sets of approximately the same size, in such a way that the learning takes place in $k - 1$ sets and the model is independently validated in the remaining set. This independence in the validation process avoids the inverse crime of using the same structure in the learning and training process. This is performed k times using all the k folds as validation sets once. Fig. 8 shows a 3-fold cross-validation. The estimate of a given parameter when resampling techniques are used is the mean of the statistics evaluated on each model over the corresponding test data. The k-fold cross-validation uses the data set more effectively and is employed in this paper to evaluated the expected risk.

[2] The risk $R(w) = \int L(d, f(x, w))dF(x, d)$ can be written using the joint probability of x and d $p(x, d)$, $R(w) = \int L(d, f(x, w))p(x, d)dx\ dd$

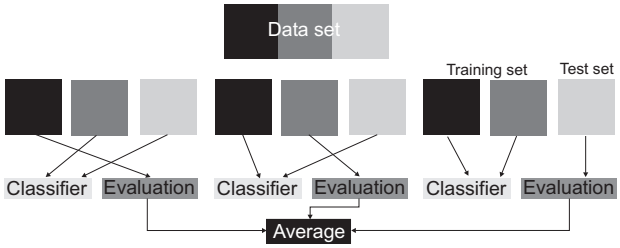

Fig. 8. This figure shows a 3-fold cross-validation. The two sets labeled as train are used to train the neural network and the error is evaluated on the test set. After all the folders are used once as test set the expected risk is estimated as the average error of each set.

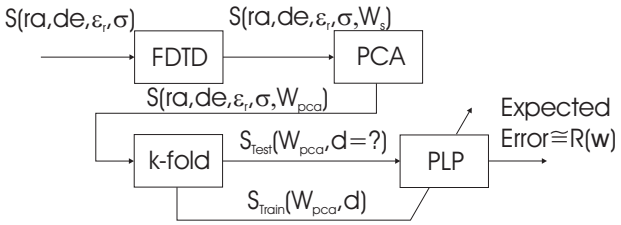

Fig. 9. This figure shows an overview of the detection system employed in this paper. Firstly, given a set S of radius (ra), depths (de), ϵ_r and σ a scattered wave ($W_s \in \mathbb{R}^p$) is calculated using FDTD. Afterwards, given $q < p$ a dimensional reduction is applied in W_s generating $W_{pca} \in \mathbb{R}^q$. Nextly, the W_{pca} is used in a k-fold system such that S_1 is used to estimate the expected error, Eq. 14, and S_{k-1} is used to adapt the PLP network parameters. In this work a 10-fold cross-validation was employed.

5.4 Neural Networks results

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

$$L(d_r) = \frac{|d_t - d_r|}{d_r}, \tag{16}$$

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). The expected value of the test risk, Eq. 14, given the loss functional defined in Eq. 16 was calculated using the 10-fold cross-validation. In Tables 3 and 5 are shown the mean deviation, $mean(L)$, the maximum deviation, $max(L)$ and the train and test times considering the dimension reduction to $q = 51, 139$ *and* 286. The system developed in this paper using a PLP network trained with the scattered wave calculated with dimensional reduction based on PCA is presented in Fig. 9. **The PLP topology is shown in Fig. 10 (10). Further details concerning the PLP topology, including its computational cost compared to**

other standard topologies, are presented in (17).

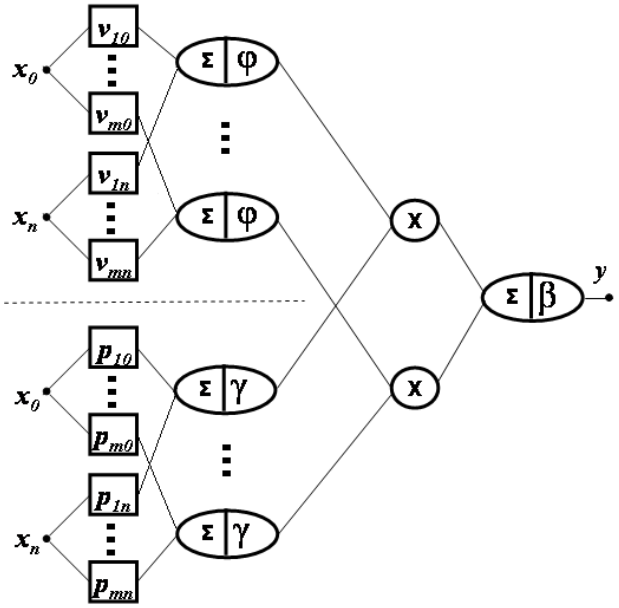

Fig. 10. The Parallel Layer Perceptron (PLP) Topology.

The PLP was trained with two parallel layers, one linear and the other with sigmoidal activation function. The PLP network was trained using a combination of the Least-Squares-Estimate (LSE) to define the linear parameters, as in (10), and the Levenberg-Marquardt algorithm (LMA) to the non-linear parameters. The learning rate was adaptively defined as in (14) and no momentum term was added. The convergence was considered given a training error bellow 0.001 **or** 15 **epochs. The number of neurons was determined using a** 3**-fold cross validation in the training set, i.e.,** 9 **folds define the training set, and a** 3**-fold cross validation was applied on them to define the number of neurons. The** 10**-th folder was used to evaluate the expected error. The total number of samples is** 1640 **and they are divide as in the folders as aforementioned.**

Table 3
Results of the Relative Error considering the *depth* prediction.

Configuration	$mean(L)$	$max(L)$	train(s)	test(s)
$PLP(286)$	0.001%	0.04%	8.44	0.016
$PLP(139)$	0.003%	0.15%	4.48	0.01
$PLP(51)$	1%	20%	6.2	0.011

Table 4 presents 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 start overfitting afterwards. The number of neurons were determined using cross-validation. However, the results given different numbers of neurons are also acceptable results, which can be used in a real-world problem.

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

Configuration	No. Neurons	$mean(L)$	$max(L)$
$PLP(286)$	6	0.009%	0.27%
$PLP(286)$	7	0.008%	0.16%
$PLP(286)$	8	0.003%	0.05%
$PLP(286)$	9	0.001%	0.04%
$PLP(286)$	10	0.004%	0.09%
$PLP(286)$	11	0.004%	0.09%
$PLP(286)$	12	0.005%	0.10%
$PLP(286)$	13	0.006%	0.15%
$PLP(286)$	14	0.005%	0.16%
$PLP(286)$	15	0.008%	0.10%

The error to each given sample is presented in Fig. 11 considering the 9 neurons configuration. All the validation data from each fold of the cross validation method is presented, i.e., the error is shown to a given sample when it is in the validation folder, not in the training one.

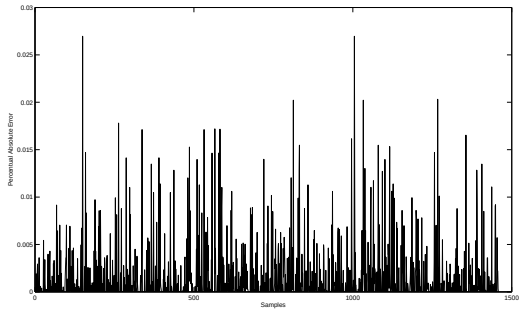

Fig. 11. This shows a characteristic Percentual Absolute Error considering the 9 Neurons configuration and the cross validation.

The results of Tables 3 and 5 show that the proposed system which combines a PCA pre-processor with a PLP is a very promising idea for assessment of inclusion in non-homogenous concrete structures.

Table 5
Results of the Relative Error for *radius* prediction.

Configuration	$mean(L)$	$max(L)$	train(s)	test(s)
$PLP(286)$	0.022%	0.94%	7.5	0.03
$PLP(139)$	1%	20%	11.1	0.018
$PLP(51)$	1%	22%	6.1	0.013

Retaining only 286 of the initial 1200 points the information is still represented without any loss, thus, 100% of the data variance was kept. The reduction to 139 and 51 dimensions resulted in 99.99% and 99% of the original variance respectively. It is clear in the results that it is harder to predict the *radii* than the *depth*. This can be physically understood by the fact that the antenna position is in front of the objects.

The non-homogeneous problem was also solved using the approach in (9). The two different configurations proposed to solve this inverse problem were a network with multiple outputs, and independent networks in parallel. The network with multiple outputs, called here $C1$, calculates simultaneously the *depth* and *radii* given as $d_1,\ d_2$ and d_3. The parallel networks, called here $C2$, calculate independently each output, given the measured variables $d_1,\ d_2$ and d_3. In $C3$ in which the *depth* is calculated using d_1, d_2 and d_3 and the *radii* are calculated in a second step (asynchronously) using also the calculated *depth*.

As all the techniques apply similar mechanism to reconstruct the *depth* their error are also very similar, as indicated in the second column of Table 6. It can

Table 6
Results considering the two configurations of the Neural Networks studied in (Caorsi *et al., 2005).*

Configuration	$Err(depth)$	$Err(radii)$
$C1$	5.8%	12.9%
$C2$	5.7%	13.4%
$C3$	5.7%	5.8%

be seen in Table 6 that an important error arises when applying this technique to a non-homogeneous problem. These results support the initiative of using a PLP network trained with the scattered wave calculated with dimensional reduction based on PCA.

Dimensional reduction was also considered in (9) but was done empirically using only three features, the peak amplitude of the reflected field; the delay of the first reflected echo, calculated with respect to the time of arrival (at the receiving point) of the direct field and a measure of the duration of the scattered field. The error average of the best configuration presented in (9) was

1.46% for the depth reconstruction which is higher than the result presented in this paper.

6 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 was defined with respect to the image reconstruction. Three $2D$ reconstruction algorithms were implemented and new configurations were proposed including the use of Artificial Neural Networks and a hybrid RTMA/PSO. 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. For all cases, the targets were successfully reconstructed with the $2D$ algorithms and the convergence was assessed in terms of the relative field error.

The main strength of the migration algorithm (non-parametric class) is its robustness to find the number of inclusions and their centers. However, it cannot identify the other physical properties that are of interest as *depth* and σ. This algorithm assumes that the background medium is known - a common assumption in many commercial softwares for the interpretation of ground penetrating radar (GPR) data (5). In addition, in this simulation line sources were used which can create a problem when the objective is to find targets in close proximity. In real problems, this algorithm can be used to locate conductors and buried pipes but it can not find the exact characteristics of the inclusions. The PSO (parametric class) can solve the general problem as the number of unknown are easily included in the formulation. Its main drawback is the computational time as it tends to require many objective function evaluations. The non-parametric algorithms are more complex to implement than the parametric ones but usually faster, as it does not require iterative solutions. This paper has proposed a hybrid parametric/non-parametric method which takes advantages of the algorithms main strengths. The hybrid approach applies a non-parametric step to identify the number of inclusions and their centers and the parametric one to identify the other parameters. The results indicated that this is better than using the method isolated.

The use of intelligent systems was also investigate in this work. Artificial Neural Networks were applied to estimate the scatter model using a training data set. It was discussed the curse of dimensionality and the use of PCA to decrease this shortcoming. Some remarkable results were presented. The ANN can be viewed as an off-line parametric model, instead of on-line as for the PSO case. For the ANN a data set, which is usually expensive to compute, is extracted from the forward model and it is used in the training (model

estimation) step. In practical terms, an equipment with ANN would required a huge off-line computational effort in the design process but little on-line effort after the scattered wave is obtained in the field (the parameters of the ANN is found before the scattered field is measured). The PSO, on the other hand, needs no off-line preparation, but the model parameters are defined after the scattered wave is obtained. Even though the PSO needs less function evaluation to find the model, as it already knows the scattered wave, it runs on-line, therefore, the operator has to wait a substantial time to get the result. The PSO is also more precise than the ANN as it can run until the objective function gets to a desired level.

Ideally, it would be interesting to combine all the three techniques to obtain a robust and fast equipment to the detection and characterization of inclusions in concrete structures. Finding the most suitable ways to combine these technique is an ongoing research.

In conclusion, significant algorithmic flexibilities utilizing 2D based algorithm were demonstrated in terms of accommodating various forms of inclusions in single and iterative reconstruction.

References

[1] C. J. Leuschen and R. G. Plumb, "A Matched-Filter-Based Reverse-Time Migration Algorithm for Ground-Penetrating Radar Data," *IEEE Trans. Geoscience and Remote Sensing*, vol. 3, pp. 295-326, 2001.

[2] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," *Proceedings of the Sixth International symposium on Micro Machine and Human Science*, pp. 39-43, 1995.

[3] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," *Proceedings of IEEE International Conference on Neural Networks*, vol. 4, pp. 1942-1948, 1995.

[4] D. J. Daniels, *Ground Penetrating Radar*, The institute of Electrical Engineers, London, 2004.

[5] *GSSI Handbook For Radar Testing of Concrete*. North Salem, NH: Published by Geophysical Survey Systems, Inc. 13, 2001.

[6] H. Jia, T. Takenaka, and T. Tanaka, "Time-Domain Inverse Scattering Method for Cross-Borehole Radar Imaging," *IEEE Trans on Geoscience and Remote Sensing*, vol. 40, no. 7, pp. 1640-1647, 2002.

[7] R. Demining and A. J. Devaney, "Diffraction tomography for multimonostatic ground penetrating radar imaging, " *Inv. Probl.*, vol. 13, pp. 29-45, 1997.

[8] S. Caorsi and G. Cevini, "Neural Networks Trained by Scattered Electromagnetic Data for GPR Applications," *2d International Workshop on Advanced GPR*, pp. 14-16, May, 2003.

[9] 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, pp. 3-7, Jan, 2005.

[10] W. M. Caminhas, D. A. G. Vieira and J. A. Vasconcelos, "Parallel Layer Perceptron," *Neurocomputing*, vol. 55, pp. 771-778-7, Oct, 2003.

[11] S. Hoole, "Artificial neural networks in the solution of inverse electromagnetic field problems," *IEEE Transactions on Magnetics* vol. 29, pp. 1931-1934, 1993.

[12] J. Gazdag and P. Sguazzero, "Migration of seismic data," *Proc. IEEE* vol. 72, pp. 1302-1315, 1984.

[13] A. J. Berkhout, "Wave field extrapolation techniques in seismic migration-A tutorial," *Geophysics* vol. 46, pp. 1638-1656, 1981.

[14] J. S. R. Jang, "ANFIS: Adaptative-Network-Based Fuzzy Inference Systems," *IEEE Transactions on Systems, Man and Cybernetics* vol. 23, pp. 665-685, May 1993.

[15] A. J.Berkhout, *Seismic Migration-Imaging of Acoustic Energy by Wave Field Extrapolation-A. Theorectical Aspects*, Elsevier, New York, 1982.

[16] Simon Haykin, *Adaptive filter theory (3rd ed.)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[17] D. A. G. Vieira, W. M. Caminhas, and J. A. Vasconcelos, "Extracting sensitivity information of electromagnetic devices models from a modified anfis topology," *IEEE Transaction on Magnetics*, vol. 40, no. 2, pp. 1180-1183, 2004.

[18] H. Ceylan, E. Tutumluer, and E. J. Barenberg, "Artificial Neural Networks for Analyzing Concrete Airfield Pavements Serving the Boeing B-777 Aircraft," *Journal of the Transportation Research Board*, Transportation Research Record 1684, pp. 110-117, 2000.

[19] R. W. Meier, *Backcalculation of Flexible Pavement Moduli from Falling Weight Deflectometer Data Using Artificial Neural Networks*. Ph.D. Dissertation, Georgia Institute of Technology, School of Civil and Environmental Engineering, Atlanta, March 1995.

[20] T. Hastie, R. Tibshirani and J. H. Friedman, *The Elements of Statistical Learning*, Springer, 2001.

[21] V. N. Vapnik, R. Tibshirani and J. H. Friedman, *Statistical Learning Theory*, New York: Wiley, 1998.

[22] V. N. Vapnik, *The Nature of Statistical Learning Theory (Statistics for Engineering and Information Science)*, Springer, 2001.