IAENG International Journal of Computer Science, 33:2, IJCS_33_2_3

Computational Investigation on the Use of FEM and RBF Neural Network in the Inverse Electromagnetic Problem of Parameter Identification

T. Hacib, M. R. Mekideche and N. Ferkha

Abstract— **In this article an attempt is made to study the applicability of a general purpose, supervised feed forward neural network with one hidden layer, namely radial basis function (RBF) neural network and finite element method (FEM) to solve the inverse problem of parameter identification. The methodology used in this study consists in the simulation of a large number of variations of magnetic relative permeability and electric conductivity in a material under test by FEM. Then the obtained results are used to generate a set of vectors for the training of a RBF neural network. Finally, the obtained neural network is used to identify a electromagnetic parameters of a group of new materials that not belonging to the original dataset. Performance of the RBF neural network was also compared with the most commonly used multilayer perceptron network model. and the results show that RBF network performs better than multilayer perceptron network model.**

Index Terms— **FEM, inverse electromagnetic problem, RBF neural network, parameter identification.**

I. Introduction

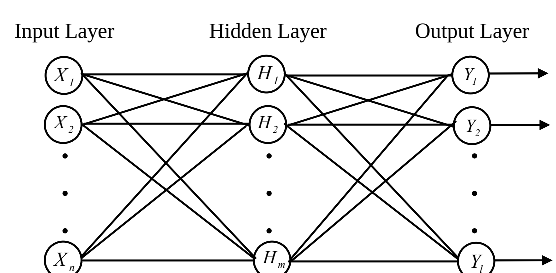
Multilayer Perceptron (MLP) network models are the popular network architectures used in most of the research applications in medicine, engineering, mathematical modelling, etc [1][2]. In MLP, the weighted sum of the inputs and bias term are passed to activation level through a transfer function to produce the output, and the units are arranged in a layered feed-forward topology called Feed Forward Neural Network (FFNN). The schematic representation of FFNN with n inputs, m hidden units and one output unit along with the bias term of the input unit and hidden unit is given in Fig. 1. An artificial neural network (ANN) has three layers: input layer, hidden layer and output layer. The hidden layer vastly increases the learning power of the MLP. The transfer or activation function of the network modifies the input to give a desired output. The transfer function is chosen such that the algorithm requires a response function with a continuous, single valued with first derivative existence. Choice of the number of the hidden layers, hidden nodes and type of activation function plays an important role in model constructions [3].

Radial basis function (RBF) neural network is based on supervised learning. RBF networks were independently proposed by many researchers [4][5] and are a popular alternative to the MLP. RBF networks are also good at modelling nonlinear data and can be trained in one stage rather than using an iterative process as in MLP and also learn the given application quickly. They are useful in solving problems where the input data are corrupted with additive noise. The transformation functions used are based on a Gaussian distribution. If the error of the network is minimized appropriately, it will produce outputs that sum to unity, which will represent a probability for the outputs. The objective of this article is to study the applicability of RBF to solve the inverse problem and compare the results with MLP.

In this paper we present an investigation on the use of FEM and ANN in the identifications of relative magnetic permeability and electric conductivity of metallic walls. The methodology consists of the following steps:

1- A large number of metallic walls with different magnetic permeability are simulated using the finite element method.
2- The obtained results are then used to generate the training vectors for artificial neural network.
3- The trained network is used to identify new parameters in the metallic wall, which not belong to the original dataset.
4- The network weights can be embedded in an electronic device, and used to identify parameters in real pieces, with characteristics similar to those of the simulated ones.

For this methodology, the measured values are independent of the relative motion between the probe and the piece under test. In other words, the movement is necessary only to change the position of the probes, to acquire the field's values, which are necessary for the identification of new parameters.



Manuscript received January 22, 2007.
T. Hacib, M. R. Mekideche and N. Ferkha are with Electrical Engineering Department, Electrical engineering study and modelling laboratory, Jijel University, 18000 Algeria (phone: +213-34 50 11 89; fax: +213-34 50 26 89; e-mail: tarik_hacib@mail.univ-jijel.dz, mek_moh@mail.univ-jijel.dz and nassira_ferkha@yahoo.fr).

(Advance online publication: 24 May 2007)

Fig. 1. Feed forward neural network

II. RBF Network Model

The RBF network has a feed forward structure consisting of a single hidden layer of J locally tuned units, which are fully interconnected to an output layer of L linear units. All hidden units simultaneously receive the n-dimensional real valued input vector X (Fig. 2). The main difference from that of MLP is the absence of hidden-layer weights. The hidden-unit outputs are not calculated using the weighted-sum mechanism/sigmoid activation; rather each hidden unit output Z_j is obtained by closeness of the input X to an n-dimensional parameter vector μ_j associated with the jth hidden unit [4].

The response characteristicsof the jth hidden unit $(j = 1, 2, \ldots, J)$ is assumed as

$$Z = K\left(\frac{\|X - \mu_j\|}{\sigma_j}\right) \tag{1}$$

where K is a strictly positive radially symmetric function (kernel) with a unique maximum at its 'centre' μ_j and which drops off rapidly to zero away from the centre. The parameter σ_j is the width of the receptive field in the input space from unit j. This implies that Z_j has an appreciable value only when the distance $\|X - \mu_j\|$ is smaller than the width σ_j. Given an input vector X, the output of the RBF network is the L-dimensional activity vector Y, whose lth component $(l = 1, 2, \ldots, L)$ is given by,

$$Y_l(X) = \sum_{j=1}^{J} w_{lj} Z_j(X) \tag{2}$$

For $l = 1$, mapping of (1) is similar to a polynomial threshold gate. However, in the RBF network, a choice is made to use radially symmetric kernels as 'hidden units'.

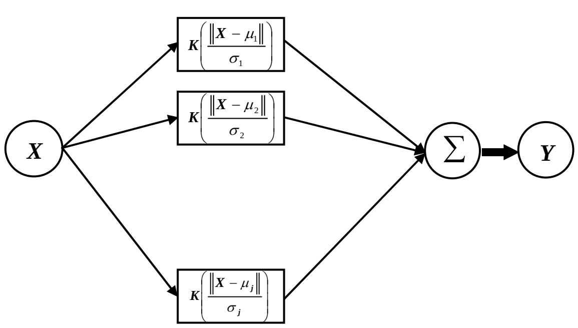


Fig. 2. Radial basis function neural network

RBF networks are best suited for approximating continuous or piecewise continuous real-valued mapping $f : R^n \rightarrow R^L$, where n is sufficiently small. These approximation problems include interpolation problems as a special case. From (1) and (2), the RBF network can be viewed as approximating a desired function $f(X)$ by superposition of non-orthogonal, bell-shaped basis functions. The degree of accuracy of these RBF networks can be controlled by three parameters: the number of basis functions used, their location and their width [3][4].

In the present work we have assumed a Gaussian basis function for the hidden units given as Z_j for $j = 1, 2, \ldots, J$, where

$$Z_j = \exp\left(-\frac{\|X - \mu_j\|^2}{2\sigma_j^2}\right) \tag{3}$$

and μ_j and σ_j are mean and the standard deviation respectively, of the jth unit receptive field and the norm is the Euclidean.

A. Training of RBF Neural Networks

A training set is an m labeled pair $\{X_i, d_i\}$ that represents associations of a given mapping or samples of a continuous multivariate function. The sum of squared error criterion function can be considered as an error function E to be minimized over the given training set. That is, to develop a training method that minimizes E by adaptively updating the free parameters of the RBF network. These parameters are the receptive field centres μ_j of the hidden layer Gaussian units, the receptive field widths σ_j, and the output layer weights w_{ij}. Because of the differentiable nature of the RBF network transfer characteristics, one of the training methods considered here was a fully supervized gradient-descent method over E [4]. In particular, μ_j, σ_j and w_{ij} are updated as follows:

$$\Delta\mu_j = -\rho_\mu \frac{\partial E}{\partial \mu_j}, \tag{4}$$

$$\Delta\sigma_j = -\rho_\sigma \frac{\partial E}{\partial \sigma_j}, \tag{5}$$

$$\Delta w_{ij} = -\rho_w \frac{\partial E}{\partial w_{lj}}, \tag{6}$$

where ρ_μ, ρ_σ and ρ_w are small positive constants. This method is capable of matching or exceeding the performance of neural networks with back-propagation algorithm, but gives training comparable with those of sigmoidal type of FFNN [6].

The training of the RBF network is radically different from the classical training of standard FFNNs. In this case, there is no changing of weights with the use of the gradient method aimed at function minimization. In RBF networks with the chosen type of radial basis function, training resolves itself into selecting the centres and dimensions of the functions and calculating the weights of the output neuron. The centre, distance scale and precise shape of the radial function are parameters of the model, all fixed if it is linear. Selection of the centres can be understood as defining the optimal number of basis functions and choosing the elements of the training set used in the solution. It was done according to the method of forward selection [4]. Heuristic operation on a given defined training set starts from an empty subset of the basis functions. Then the empty subset is filled with succeeding basis functions with their centres marked by the location of elements of the training set; which generally decreases the sum-squared error or the cost function. In this way, a model of the network constructed each time is being completed by the best element. Construction of the network is continued till the criterion demonstrating the quality of the model is fulfilled. The most commonly used method for estimating generalization error is the cross validation error.

III. Electromagnetic Field Computation

In this study, the magnetic field is calculated using the finite element method. This method is based on the $\boldsymbol{A}$ representation of the magnetic field. The calculations are performed in two steps. First, the magnetic field intensity is calculated by solving the system of equations:

$$\boldsymbol{rot}\left(\boldsymbol{E}\right)=-\frac{\partial \boldsymbol{B}}{\partial t} \tag{7}$$

$$\boldsymbol{rot}\left(\boldsymbol{H}\right)=\boldsymbol{J} \tag{8}$$

$$div\left(\boldsymbol{B}\right)=0 \tag{9}$$

where $\boldsymbol{H}$ and $\boldsymbol{E}$ are the magnetic and electric field respectively, $\boldsymbol{B}$ the magnetic induction and $\boldsymbol{J}$ the electric current density. This system of equations is coupled with relations associated to material property, material being assumed to be isotropic:

$$\boldsymbol{B}=\mu\left(\left|\boldsymbol{H}\right|\right)\boldsymbol{H} \tag{10}$$

$$\boldsymbol{J}=\sigma \boldsymbol{E} \tag{11}$$

where μ is the magnetic permeability, σ is the electrical conductivity.

The magnetic vector potential $\boldsymbol{A}$ is expressed by:

$$\boldsymbol{B}=\boldsymbol{rot}\left(\boldsymbol{A}\right) \tag{12}$$

The electromagnetic field analysis for a cartesian system is carried out by the finite element method. The equation of the electromagnetic field is expressed by the magnetic vector potential $\boldsymbol{A}$:

$$\boldsymbol{rot}\left(\frac{1}{\mu}\boldsymbol{rot}\,\boldsymbol{A}\right)+\sigma\frac{\partial \boldsymbol{A}}{\partial t}=\boldsymbol{J}_s \tag{13}$$

where $\boldsymbol{J}_s$ is the vector of supply current

Equation (13) is discretized using the Galerkin finite element method, which leads to the following algebraic matrix equation:

$$\left(\left[\boldsymbol{K}\right]+j\omega\left[\boldsymbol{C}\right]\right)\left[\boldsymbol{A}\right]=\left[\boldsymbol{F}\right] \tag{14}$$

$$\boldsymbol{A}=\sum \alpha_j\left(x,y\right)\boldsymbol{A}_j \tag{15}$$

α_j is the interpolation function.

$$\boldsymbol{K}_{ij}=\iint_{\Omega}\frac{1}{\mu}\,\boldsymbol{grad}\,\alpha_i\;\boldsymbol{grad}\,\alpha_j\;dxdy \tag{16}$$

$$\boldsymbol{C}_{ij}=\iint_{\Omega}\sigma\,\alpha_i\alpha_j\;dxdy \tag{17}$$

$$\boldsymbol{F}_i=\iint_{\Omega}\boldsymbol{J}_s\,\alpha_i\;dxdy \tag{18}$$

α_i is the projection function.

In the second step, the field solution is used to calculate the magnetic induction $\boldsymbol{B}$.

More details about the finite element theory can be found in [7].

IV. Methodology For Parameter Identification

First of all, an electromagnetic device was idealized to be used as an electromagnetic field exciter (fig. 3). In this paper, we have considered direct current in the coils. So, the material of the metallic wall must be ferromagnetic. To increase the sensitivity of the electromagnetic device a magnetic core with a high permeability is used and the air gap between the core and the metallic wall is reduced to a minimum.

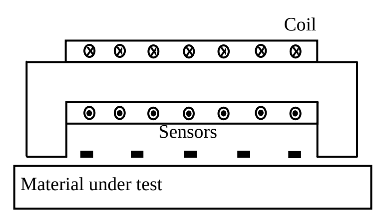


Fig. 3. Arrangement for the measurements

Deviations of the magnetic induction (difference in magnetic induction without and with material under test) at equally stepped points in the external surface of the material under test are taken.

Fig. 4 shows the steps of the methodology used in this work. Steps 1-4 correspond to the finite element analysis.

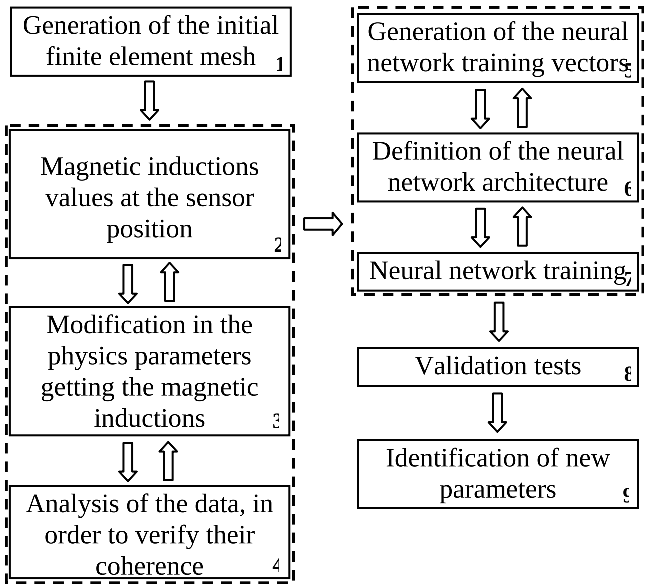


Fig. 4. Flowchart of the used methodology

The problem was solved under Matlab® workspace using the partial differential equation toolbox and neural network toolbox for the mesh generation and neural networks architecture definition. For the finite elements problem and the inverse problem solution, we use programs developed by us.

The simulations were done for a hypothetic metallic wall with 1mm height and 15mm width. The material of the wall is a magnetic material. The relative permeability of the core is supposed to be 500 and the air gap is 0.1mm. Finite element meshes with 36000 elements and 18000 nodes, approximately, were used in the simulations. Fig. 5 shows a field distribution for one of these simulations.

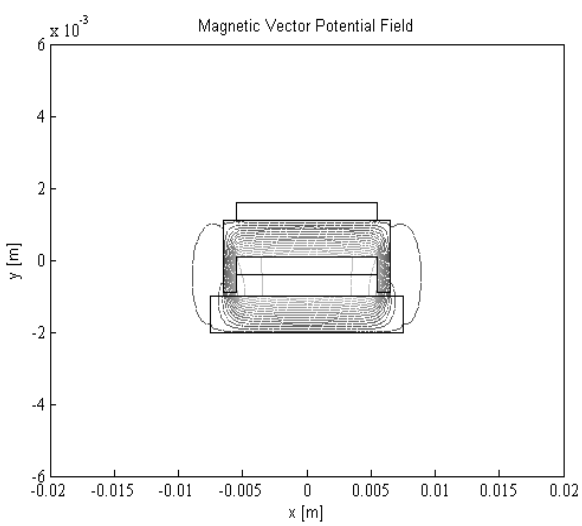


Fig. 5. Solution in magnetic potential vector ***A***

Fig. 6 and 7 shows the evolution of the magnetic induction in the region of the device at the sensor position without and with metallic wall (material under test) respectively.

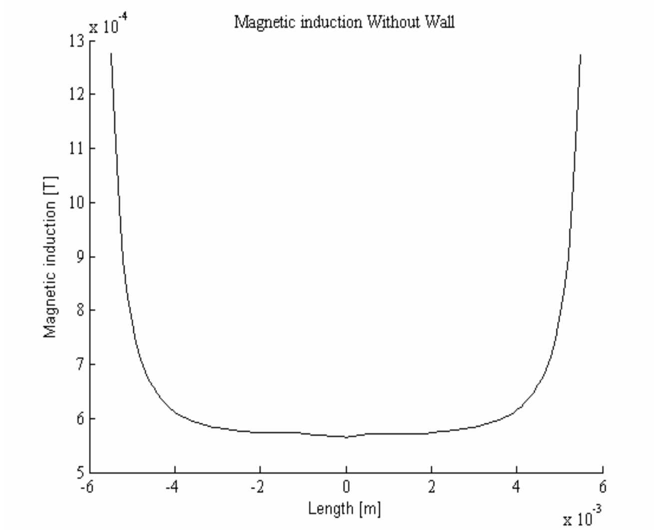


Fig. 6. Magnetic induction field without metallic wall

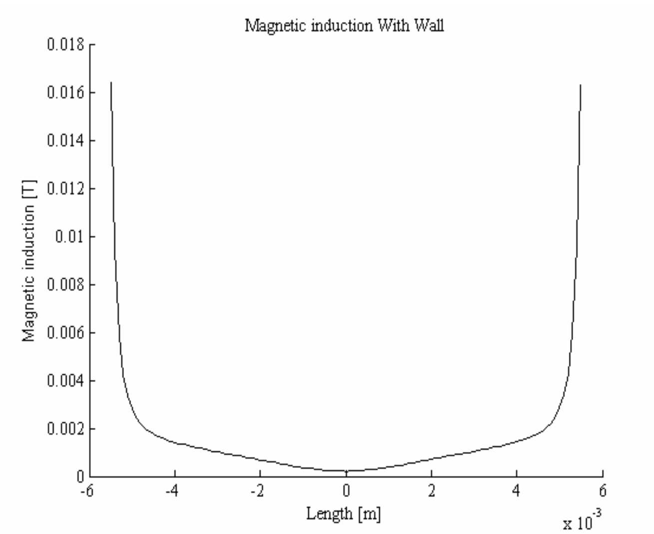


Fig. 7. Magnetic induction field with metallic wall

During the phase of finite elements simulations, errors can appears, due to its massively nature. So, the results of the simulations must be carefully analyzed. This can be done, for instance, plotting in the same graphic the magnetic induction deviations for a set of parameters. Fig. 8 shows the magnetic induction deviation in the region of the device at the sensor position for three materials having the same electrical conductivity (10^3 $[(\Omega/m)^{-1}]$), and relative magnetic permeability ranging from 50 to 300. A similar graphic, with electrical conductivity equal to 10^6 $[(\Omega/m)^{-1}]$ and magnetic relative permeability ranging from 50 to 300 is shown in Fig. 9. Fig 10 shows the graphics for a fixed magnetic relative permeability (240), and three different electrical conductivity ranging from 6 10^5 $[(\Omega/m)^{-1}]$ to 10^8 $[(\Omega/m)^{-1}]$. Fig. 11 shows a similar graphic, for the magnetic relative permeability equal to 520. In this graphics the magnetic inductions deviations are at vertical axes and length are at horizontal axes.

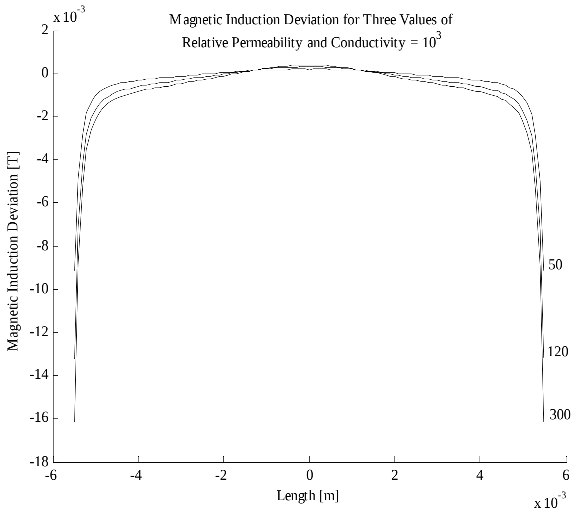


Fig. 8. Magnetic induction deviation for three values of magnetic relative permeability and electrical conductivity equal to 10^3

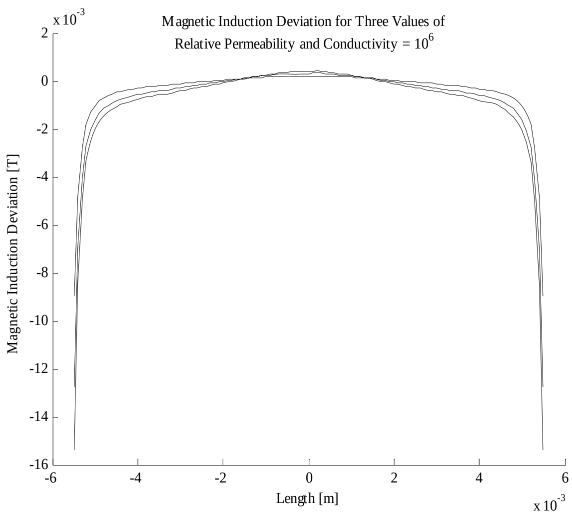


Fig. 9. Magnetic induction deviation for three values of magnetic relative permeability and electrical conductivity equal to 10^6

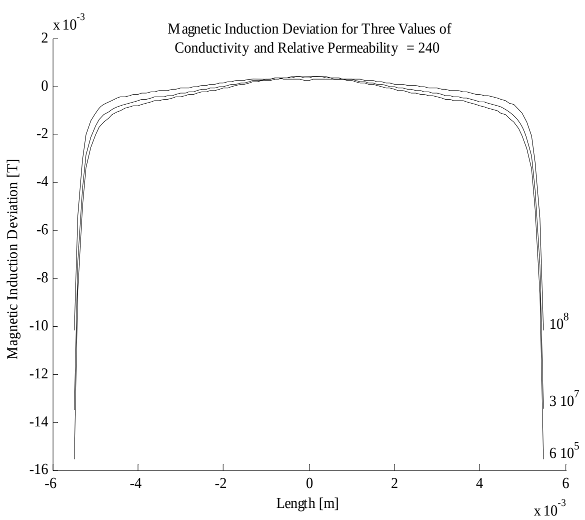


Fig. 10. Magnetic induction deviation for three values of electrical conductivity and magnetic relative permeability equal to 240

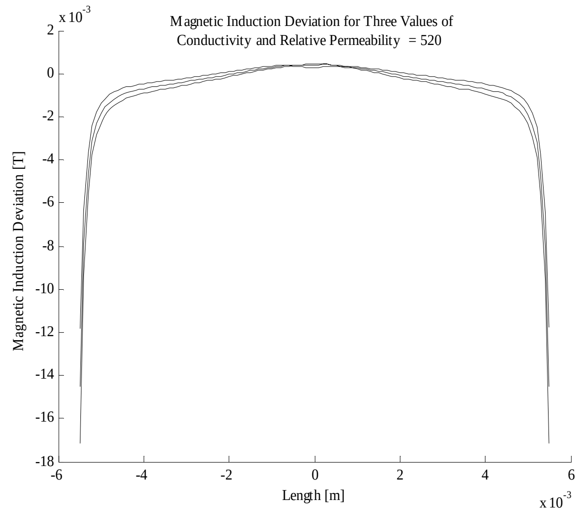


Fig. 11. Magnetic induction deviation for three values of electrical conductivity and magnetic relative permeability equal to 520

V. Formulation Of Network Models For Parameters Identifications

In the step 5, we generate the training vectors for neural network. In this work, we generated 300 vectors with 11 elements each one for the training of RBF neural network. From the original 300 vectors, 225 (75 %) were used in the network training, and 75 (25 %) were used in their validation.

The RBF neural network architecture considered for this application was a single hidden layer with Gaussian RBF. The basis function ϕ is a real function of the distance (radius) r from the origin, and the centre is c. The most common choice of ϕ includes thin-plate spline, Gaussian and multiquadric. Gaussian-type RBF was chosen here due to its similarity with the Euclidean distance and also since it gives better smoothing and interpolation properties [5]. The choice of nonlinear function is not usually a major factor in network performance, unless there is an inherent special symmetry in the problem.

Training of the RBF neural network involved two critical processes. First, the centres of each of the J Gaussian basis functions were fixed to represent the density function of the input space using a dynamic 'k means clustering algorithm' [8]. This was accomplished by first initializing the set of Gaussian centres μ_j to random values. Then, for any arbitrary input vector $X^{(t)}$ in the training set, the closest Gaussian centre, μ_j is modified as:

$$\mu_j^{new} = \mu_j^{old} + \alpha\left(X^{(t)} - \mu_j^{old}\right) \tag{19}$$

where α is a learning rate that decreases over time. This phase of RBF network training places the weights of the radial basis function units in only those regions of the input space where significant data are present. The parameter σ_j is set for each Gaussian unit to equal the average distance to the two closest neighbouring Gaussian basis units. If μ_1 and μ_2 represent the two closest weight centres to Gaussian unit j, the intention was to size this parameter so that there were no gaps between basis functions and only minimal overlap between adjacent basis functions were allowed. After the Gaussian basis centres were fixed, the second step of the RBF network training process was to determine the weight vector W which would best approximate the limited sample data X , thus leading to a linear optimization problem that could be solved by ordinary least squares method. This avoids the problem of gradient descent methods and local minima characteristic of back propagation algorithm [9].

For MLP network architecture, a single hidden layer with sigmoid activation function, which is optimal for the dichotomous outcome, is chosen. A back propagation algorithm based on conjugate gradient optimization technique was used to model MLP for the above data [10].

Fig. 12 shows the performance of a training session, and table I show some results for the validation of the network.

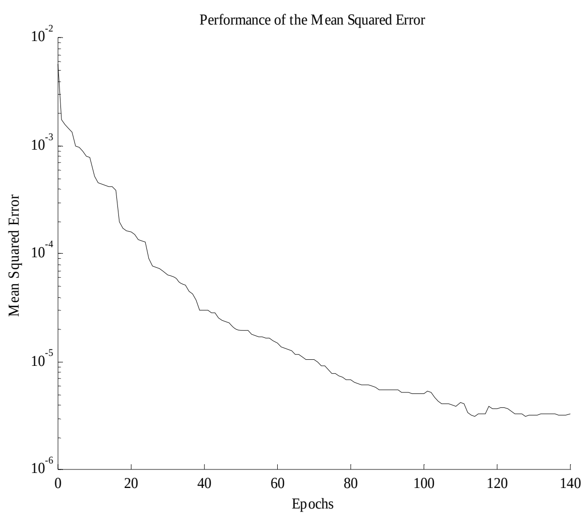


Fig. 12. Performance of the RBFNN during a training session

Table I. Expected and obtained values during a training session

Relative magnetic permeability			Electric conductivity		
Expected	RBF	MLP	Expected	RBF	MLP
287.55	287.54	287.53	1.799 e1	1.801 e1	1.802 e1
311.34	311.35	311.35	4.500 e2	4.506 e2	4.519 e2
446.21	446.21	446.24	2.250 e1	2.249 e1	2.246 e1
527.07	527.04	527.33	3.010 e5	3.011 e5	3.014 e5

VI. New Identifications

After the NN training and validations, new parameters were simulated by the FEM, for posteriori identification by the network. Table II shows the values of parameters of the material under test, and the obtained values, by the NN.

Table II. Simulation results for new parameters

Parameter	Relative magnetic permeability			Electric conductivity		
	Expected	Obtained		Expected	Obtained	
		RBF	MLP		RBF	MLP
1	89	89.001	88.933	6.500 e1	6.502 e1	6.503 e1
2	212	211.99	211.98	2.500 e2	2.501 e2	2.502 e2
3	360	360.02	359.95	2.200 e6	2.199 e6	2.201 e6
4	472	472.01	472.03	4.100 e5	4.099 e5	4.101 e5

As we can see, the results obtained in the identification of new parameters, obtained by the neural network agree very well with the expected ones, demonstrating that the association of the finite element method and RBF neural network in very powerful in the solution of inverse problems, like parameters identifications in metallic walls.

VII. Conclusion

In this paper we presented an investigation on the use of the finite element method and artificial neural network for the identification of parameters in metallic walls, present in industrial plants.

The proposed approach was found to be highly effective in identification of parameters in electromagnetic devices. Comparison of the result indicates that RBF neural network is trained and identifies the electromagnetic parameters faster than MLP neural network. Future works are intended to be done in this field, such as the use of more realistic finite element problems, computer parallel programming, in order to get quickly solutions.

References

[1] E. Coccorese, R. Martone and F.C. Morabito, “A neural network approach for the solution of electric and magnetic inverse problems", *IEEE Transactions on Magnetics, vol. 30, n°. 5, pp. 2829-2839, 1994.*

[2] N.P. De Alcantara, J. Alexandre, M. De Carvalho, “Computational investigation on the use of FEM and ANN in the non-destructive analysis of metallic tubes", *10th Biennial conference on electromagnetic field computation ICEFC*, Perugia, 2002, Italy.

[3] S. Haykin, *Neural networks: a comprehensive course*, USA, New York, 1999.

[4] P. Ramuhalli, L. Udpa and S.S. Udpa, “Electromagnetic NDE signal inversion by function-approximation neural networks", *IEEE Transactions on Magnetics, vol. 38, n°. 6, pp. 3633-3642, 2002.*

[5] T. Chady, M. Enokizono, R. Sikora, T. Todaka and Y. Tsuchida, “Natural crack recognition using inverse neural model and multi-frequency eddy current method", *IEEE Transactions on Magnetics, vol. 37, n°. 4, pp. 2797-2799, 2001.*

[6] D. Cherubini, A. Fanni, A. Montisci, P. Testoni, “A fast algorithm for inversion of MLP networks in design problems", *COMPEL the International Journal for Computation and Mathematics in Electrical and Electronic Engineering, vol. 24, n°. 3, pp. 906-920, 2005.*

[7] P.P. Silvester and R.L. Ferrari, *Finite elements for electrical engineers*, Cambridge, U.K, 1996.

[8] J. A. Hartigan, *Clustering algorithms*, John Wiley and Sons, New York, 1975.

[9] L. Mokhnache, A. Boubakeur, “Comparison of different back-propagation algorithms used in the diagnosis of transformer oil" *IEEE Conference on electrical insulation and diagnostic phenomena, pp. 244-247, 2002.*

[10] T. Hacib, M.R. Mekideche, N. Ferkha, “Defect identification using artificial neural networks and finite element method", *1st IEEE International Conference ICELIE, Hammamet, 2006, Tunisia.*

T. HACIB was born in Jijel, Algeria in 1977. he recieved the Licence degree in Electrical Engineering in 2001, and the Magister in Electrical Engineering in 2004 from the Electrical Engineering Institute of Jijel University, Algeria. He is currently working toward the Ph.D degree at Jijel University. His principal research interests include inverse problems, computations of electromagnetic field and neural networks applications in identification of parameters and defects of electromagnetic systems. Mr HACIB has participated in many international conferences and he his published in more than 10 conferences proceedings journals. He is member of LAMEL laboratory.

M. R. MEKIDECHE was born in Jijel, Algeria in 1957. After his first diploma as electrical engineer at the USTO University at Oran, he joined, in 1981, the National Polytechnicas School at Algiers as an assistant and post-graduate student preparing a Magister. He got the "Maître assistant" degree in 1986. After his PhD degree from the University of Nantes (France) he became "Maître de Conferences" and then Professor. He is now at the University of Jijel, as a teacher and Dean of the engineering faculty.
Prof. MEKIDECHE is now heading the LAMEL laboratory of the University of Jijel and is the author or co-author of about 50 works published or presented at international conferences.

N. FERKHA was born in Jijel, Algeria in 1977. She recieved the engineer degree in Electrical Engineering in 2000, and the Magister in Electrical Engineering in 2004 from the Electrical Engineering Institute of Jijel University, Algeria, where she has now an Ph.D student. Here current research interests include vibratory phenomena, computations of electromagnetic field and neural networks applications in shape optimization of electromagnetic systems. She is member of LAMEL laboratory.