Journal of Engineering and Applied Sciences 2 (4): 775-782, 2007

Tomographic Velocity Images by Artificial Neural Networks

N. Djarfour, J. Ferahtia and K. Baddari
Laboratoire de Physique de la Terre (LABOPHYT), Université de Boumerdès, Algérie

Abstract: The present study deals with the use of Elman artificial neural network (feedback connexion) to reconstruct the velocity image from a traveltime in the seismic tomography experiment. This recurrent connection provides the advantage to store values from the previous time step, which can be used in the actual time step. The backpropagation algorithm has been used to learn the suggested neural network. Efficiency of these networks has been tested in training and generalization phases. A comparative reconstruction with two classical methods was performed using backprojection and Algebraic Reconstruction Techniques (ART). The obtained results clearly show improvements of the quality of the reconstruction obtained by artificial neural networks.

Key words: Elman neuron networks training, back-propagation, traveltime, velocity, tomography, backprojection, ART

INTRODUCTION

Tomography is the method to determine the internal structure of an object starting from a set of observations that in some way carry information of the inside structure of this object. For this reason, the tomography is of great geological interest for the comprehension of the geodynamic phenomena, recognition of geological facies and the estimate of the acoustics' parameters from the geophysical data. Many research efforts in seismic tomography have been directed towards the definition of efficient and fast tools of tomographic reconstruction of velocity's image. From the mathematical point of view, this problem is a severely ill-posed inverse problem (due to high degree of non-linearity between seismic data and velocity model, sensitivity to noise in measurements) (Bioshop *et al*., 1985).

The proposed method in this study is an attempt to reconstruct the velocities' image using artificial neural networks, which have the remarkable ability to derive meaning from complicated or imprecise data. This network learns characteristics from traveltime and image as it processed them in the training phase. The goal of our proposed method is to improve the tool of tomographic reconstruction in one part and to reduce the time necessary to reach this reconstruction in the other part.

Computer simulations have been performed in order to assess the proposed technique. The results have been compared with that obtained by two conventional methods, known as backprojection and Algebraic Reconstruction Technique (ART).

Artificial neural network principles: An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The most basic element of the human brain is a specific type of cell, which provides us with the ability to remember, think and apply previous experiences to our every action. These cells are known as neurons. The power of the brain comes from the number of these basic components and the multiple connections between them. The key element of this paradigm is the novel structure of the information processing system Fig. 1a. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. Each neuron is linked to certain of its neighbours with varying coefficients of connectivity that represent the strengths of these connections. Each neuron transforms all signals that it receives in an output signal, which is communicated to other neurons. For example, the artificial neuron i multiplies each input signal by the synaptic coefficient wi and adds all these weighted entries in order to obtain a total simulation. Using an activation function (or transfer), it calculates its activity at the output, which is communicated to the following neurons Fig. 1b.

Besides their scientific value for neuroscientists, the ANN models that inspired this research have been especially useful in solving problem of speech and character recognition, classification and combinatorial optimization (Calderôn-Macias *et al*., 2000; de Groof, 1992). All these tasks for which the neural network can offer a good solution, in particular where the data are

Corresponding Author: N. Djarfour, Laboratoire de Physique de la Terre (LABOPHYT), Université de Boumerdès, Algérie

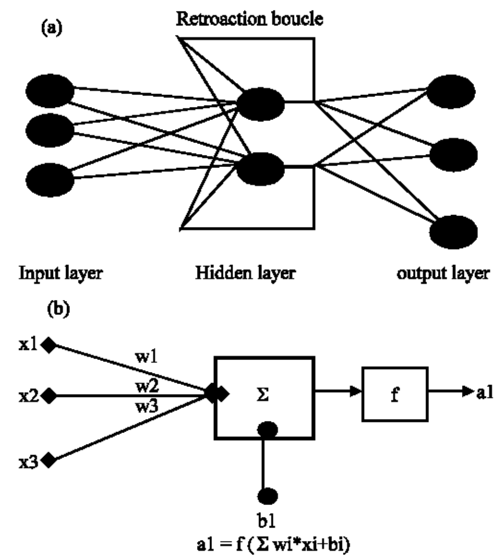

Fig. 1: Neural networks principe

noisy, or the explicit knowledge of task is not available or when unknown non-linearity between input and output may exist.

There are two phases in neural information processing. They are the learning and the generalisation Phase; In the training phase, a training data set is used to determine the weight parameters that define the neural model. In the generalisation one the neural network is used as a classical function.

Artificial neural network's training: The most widely used training method is known as backpropagation method. This later produces a least square fit between the actual network output and a desired output by computing a local gradient in terms of the network weights. The design of neural network models capable of training originates in the work of the neurophysiologist. Let us start from this simple and known thing of all, the neurons are connected the ones to the others by synapses. Hebb made the assumption that the force of a synapse increases when the neurons which it connects act in the same way at the same time. Conversely, it decreases when the neurons have different activities at the same time. Consequently the intensity of this force varies, more or less, according to the simultaneous activity of each inter-connected neurons. A learned form will correspond to a state of the network in which certain neurons will be active while others will not be. It is the same principle which is applied to artificial neurons training. All the neurons that are connected to each other, constitute a network. When one presents at the network a form to be learned, the neurons start simultaneously a state of activity which causes a small modification of the synaptic forces. It follows a quantitative reconfiguration of the whole of the synapses, some of them become very strong, the others become weak. The learned form is not directly memorized at a precise place; it corresponds in a particular energy state of the network, a particular configuration of the activity of each neuron, in a very large case of possible configurations. This configuration is supported by the values of the synaptic forces (Parisi *et al*., 1996).

Let y^s_j represents the output of the jth neuron at layer s, w^s_{ij} is the weight connecting the ith neuron in layer s to the jth neuron at layer s-l and o^s_j is the weighted sum at he input of the jth neuron of layer s. This output can be expressed by the following equation:

$$y_j^s = f(o_j^s) = f(\sum w_{ji}^s . y_i^{s-1})$$

This relation allows, by knowing the input of the first layer of the network, to gradually calculate the value of global output of the network, thus ensuring the forward propagation. When one compares this output with the desired output, one can calculate the error function, generally given by:

$$e = \frac{1}{2}(y - \bar{y})^2$$

where y is the desired output and $\bar{y}$ the obtained output. In backpropagation method, the direction in which weights are updated is given by the negative of the gradient of e with respects to every elements of the weight Note that weight updates start in the output layer and processed towards the first layer, hence the term backpropagation (Calderôn-Macias *et al*., 2000). This algorithm consists in minimizing e by a gradient descent. With each presentation of an example, one to modify the synaptic weights so as to reduce e. the modification of the weights of the connections of the neurons of the hidden layers s, with the step of calculation n, is carried out using the following relation:

$$\Delta w_{ij}^s = -\mu.(e_j^s.y_i^{s-1})_n + (\Delta w_{ji}^s)_{n-1}$$

where μ is known as the learning rate parameter and is usually a small number, say between 0 and 0.5. The quantity e^s_j is the locale error of jth neuron in the layer s defined as:

$$e_j^s = f'(O_j^s).\sum_k e_k^{s+1}.w_{kj}^{s+1}$$

Weights and bias terms are first initialized to random values. In general, there are no strict rules to determine the network configuration for optimum training and prediction.

Basic concepts of tomography: The seismic tomography uses the measured traveltimes as input to calculate the velocity distribution of the subsurface structure as output. Accordingly, the two dimensional image of the velocity variations can be obtained. The traveltime of a ray in a continuous velocity medium is given by this relation:

$$T = \int P(x)ds,$$

where T is the traveltime, while P represents slowness. The relation between the parameters and the traveltime is non-linear, since the integration path x(s) depends on the velocity. This inherent non-linearity means that the inverse problem can be very difficult to solve.

Various authors (Becht *et al*., 2004; Robert, 1992; Pratt and Worthington, 1988) have described how the region between two wells can be divided into a number of cells of constant slowness. The seismic traveltime ray then, can be approximated by the set of linear equations of the matrix form, in which. The traveltime of a particular ray is given by the sum of elementary times of the sets of the pixels between the source and the receiver. This time is given by the following formula (Jiang and Wang, 2003; Robert, 1992; Peterson *et al*., 1985):

$$T = AP,$$

where T is a vector of traveltime, the matrix A contains the distance travelled by each ray in each cell. This matrix characterizes completely the geometry of acquisition. Every line of this matrix represents the equation of a ray and the vector P is the vector to be determined contains the slowness in each cell. The reconstruction tomographic becomes a problem of solution of a set linear equations. There are basically two major categories of image reconstruction algorithms: analytic and algebraic methods. The analytic methods include e.g., the Fourier reconstruction and filtered back projection or inverse Radon transform, these methods consider the object to be imaged, described by continuous function, with continuous set of projections. Most common algebraic reconstruction methods are the ART (Algebraic Reconstruction Technique) or SIRT (Simultaneous Iterative Reconstruction Method) and matrix inversion, these method assume that the object is composed of a discrete set of number pixel, the values of pixel will be inferred from a finite number of projects (Robert, 1992). These iterative procedures modify an initial velocity model to find a best-fit solution in a least-squares sense. The full description of these methods can be found in (Robert, 1992; Kissling *et al*., 2001; Ali *et al*., 2004).

RESULTS AND DISCUSSION

We have chosen the Feedback type neuronal structure, well-known as Elman network that was firstly introduced by Professor Jeff Elman illustrated in Fig. 1a (Elman, 1990). It is an artificial neural network characterized by feedback connection from the hidden layer output to its input. This feedback loop allows Elman networks learning to recognize and generate temporal patterns, as well as spatial patterns. This internal looping ensures a recirculation of information, in fact, the delay in this connection stores values from the previous time step, which can be used in the current time step. By analogy with human brain thinking, this delayed information leads to hesitations before making a final decision by a human being.

The activation functions of the hidden layer are of logsigmoid type given by the following equation:

$$f(x) = \frac{1 - e^{-2\alpha x}}{1 + e^{+2\alpha x}}$$

where α is a parameter which controls the steepness of the function near x = 0.

Using this activation function for all hidden layer neurons, the network will be considered as a sigmoid function-based decomposition of the input signal. In order to implement the ANN to reconstruct seismic image, two velocity models will be used. The first is used to train the network, whereas the second model is used to test the network ability to reconstruct the new image using the derived weights. This latter is reconstructed by traditional methods namely ART and Backprojection. This allows us to make a comparison between neural networks and classical methods in generalisation stage.

Through out all experiments, the input of the ANN is traveltime, the raypath configuration of this traveltime is given in the Fig. 2, the output is the distribution of the parameter to be estimated (velocity) example Fig. 3. ANN with different number of neurons in the hidden layer was trained using the backpropagation method.

In order to observe the behavior of the network in the training process, we have limited the number of iterations to 25 as illustrated in the Fig. 4. According to the obtained results, the training performance has not reached the required accuracy, thus number of iterations must be

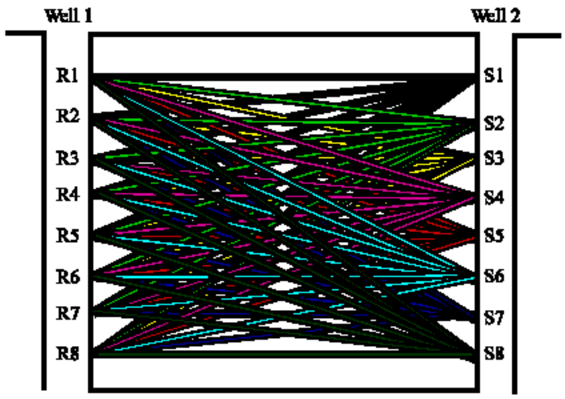

Fig. 2: Seismic data crosswell geometry

increased. To testify the obtained trained network, we have used the training traveltimes to feed the trained network and the output Fig. 5 is compared to the exact image Fig. 3. The difference between the two images, determines efficiency of the obtained neural network.

We have increased the iteration's number to 4000, we have obtained a better and precise performance as mentioned by the Fig. 6. Then we have tested the efficiency of network on training data, according to the obtained performance, we presented it the traveltime in input and the output of network is given by the Fig. 7, which is very similar to the exact image to be reconstructed given by the Fig. 3, this allows us to say that the network has been adapted to reconstruct the image velocity from seismic traveltime. A summary table of these experiments has been given in the end of this study.

In order to see the generalization's possibility of ANN, we tested the network by data calculated on the model given by Fig. 8, after have achieved the training on the model of the Fig. 3. In this model we have introduced the same complexity order than one introduced into the model used in training, which is characterized by vertical and horizontal progressive variations of the parameter to be estimated (velocity). It is this type of variation which causes a great difficulty during its reconstruction.

By comparing the reconstructed image by the network Fig. 9 with the exact one of Fig. 8, we notice that a satisfactory reconstruction has been achieved by the network in generalization phase.

In order to compare the result of application of ANN, we propose to use two classical methods. The first reconstruction has been done by backprojection method, the type of calculation of this method leads to a relatively smooth solution of the distribution of velocity image, that constitute, the average estimate of this distribution. The

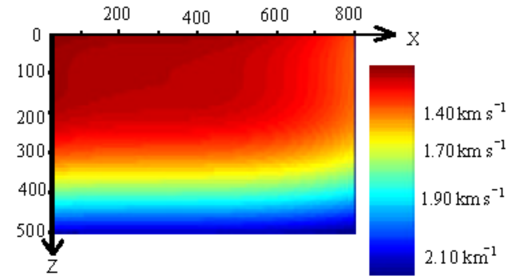

Fig. 3: Model proposed for training of ANN

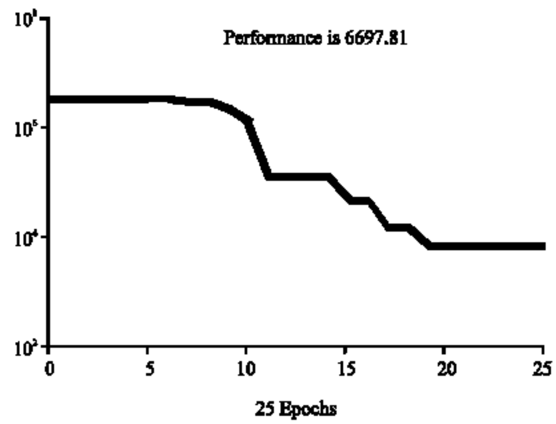

Fig. 4: Mean squared error

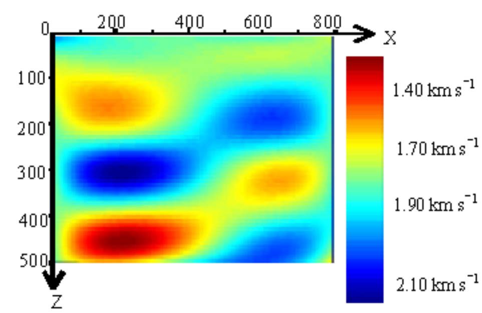

Fig. 5: Output of ANN during the training

key disadvantage of backprojection is that it cannot reconstruct image uniformly and symmetrically as shown in the Fig. 10, the result is severely distorted image compared with the original image.

In order to compare the efficiency of reconstruction by ANN, we propose to use the ART as a conventional method. In this method, it is common practice to employ a uniform image as starting model and to use the traveltime to reconstruct the image by ART algorithm. Following this conventional approach, the image is initialized with a constant value of velocity for all the

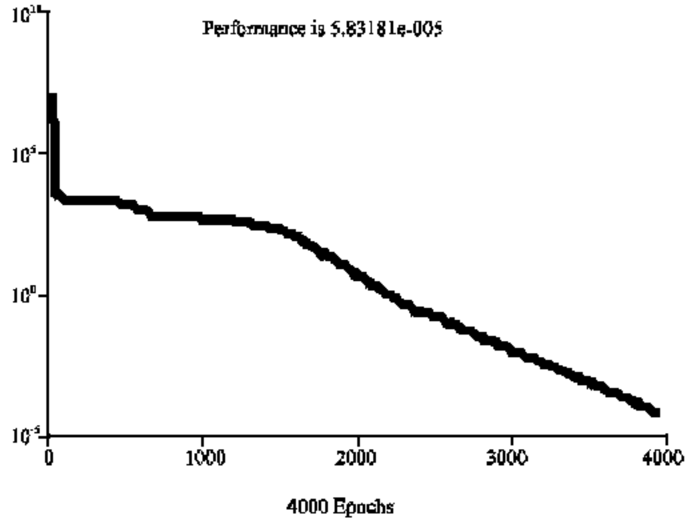

Fig. 6: Mean squared erro

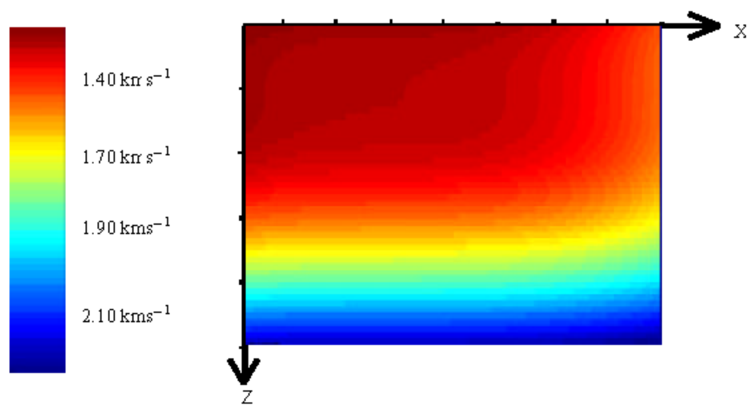

Fig. 7: Output of ANN after the training

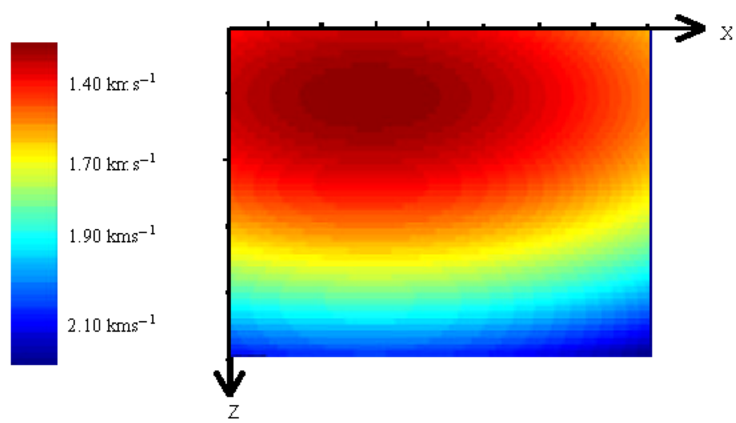

Fig. 8: Model proposed for testing the ANN in generalization phase

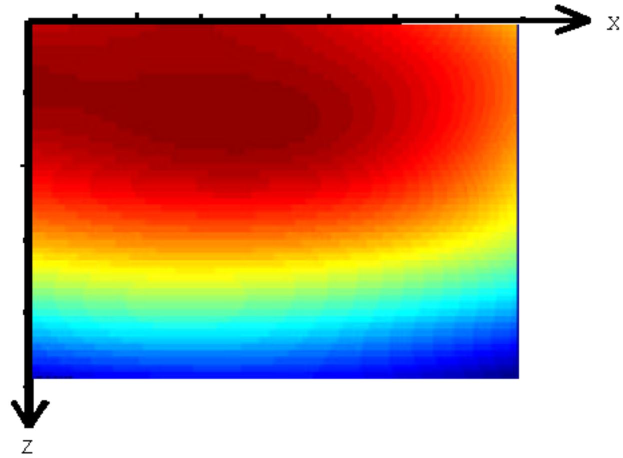

Fig. 9: Output of the ANN in generalization

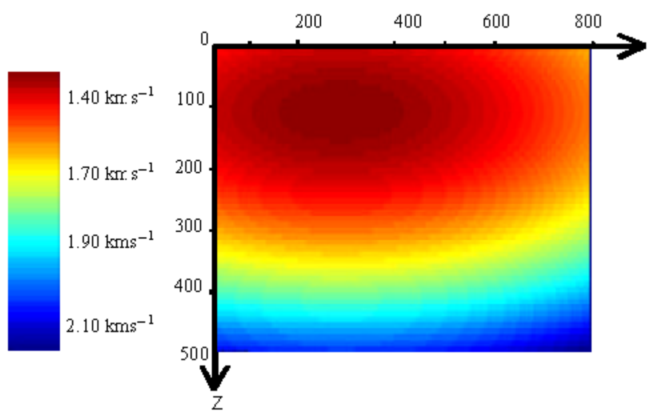

Fig. 10a: Model proposed for comparison with ANN

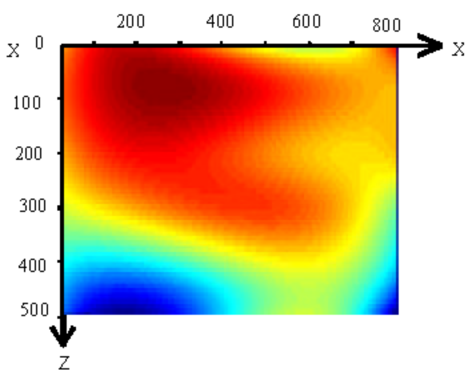

Fig. 10b: Result of backprojection

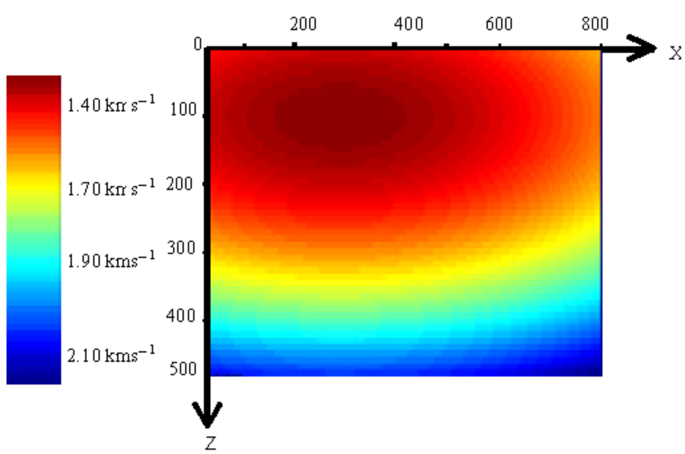

Fig. 11a: Model proposed for comparison with ANN

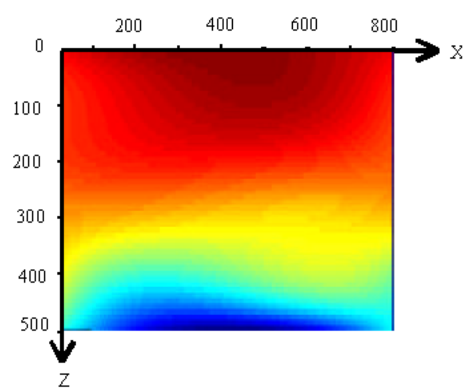

Fig. 11b: Result of ART

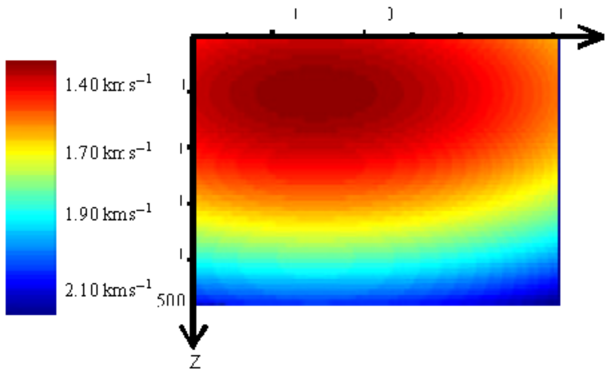

Fig. 12a: Model proposed for comparison with ANN

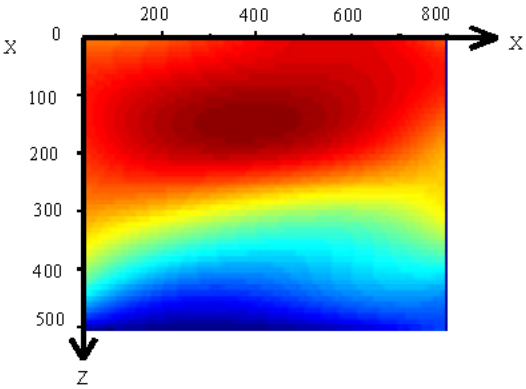

Fig. 12b: Result of ART

ANN	Neurones nembre in layer	Nembre of iterations	Performance
Hidden layer1-hidden layer 2	20-0	25	6697.8157
	20-0	4000	5.95 e-^{05}
	25-0	4000	5.83 e-^{05}
	30-0	3000	0.00916977
	5-5	4000	1889.96
Activation function hidden layer.	Log-sigmoid		
Activation function output layer	Linear		
Training algorithm	Back propagation conjugate gradient		

model, 2000 m s^{-1}. One can notice that after 10 iterations the original image could not be reconstructed sufficiently Fig. 11 and the images is contaminated with several artefacts. To improve ART reconstruction, we tested a reconstruction strategy by using the result of backprojection as initial image. We obtained the image shown in the Fig. 12 after five iterations, that constitute a better image then the one obtained in fist step (stating from uniform image), These reconstruction method depends on the specific choice of the initial values. Consequently, the effectiveness of the tomography with ART algorithm will be strongly affected when the initial image is not a good approximation of the exact image (unknown).

By comparing the result of ANN in the generalization phase with the one obtained by ART combined with backprojection, it is clear seen that the ANN has the superiority of reconstruction in quality of the image and in time necessary to obtain this reconstruction compared with the ART method, that produced a poor image reconstruction. Several regions inside the image have been altered by this reconstruction. In addition, in this method, the starting model is required and new ray paths are calculated at each iteration, so that it is large computational time, but with the ANN we get the reconstruction in only one passage, because the training is done just once and the obtained ANN can be used as classical function.

CONCLUSION

In this research the artificial neural networks were driven for the evaluation of velocity image from seismic traveltime. Results show that ANN can establish relation between the traveltime and the velocity image, by their capacity of approximation and adaptation. The training of the network can be considered as useful tool to synthesize automatically a generally function non-liner (control mapping). The network of feedback type offers the advantage of a fast and simple training while using an algorithm of back propagation conjugate gradient of the error. The efficiency of the network is tested in comparison with other techniques (ART and backprojection) and results were extensively satisfactory, in the same way possibilities of generalisation of their use, The ANN reconstructs the image more rapidly than the others traditional algorithms, because the training makes only one time.

REFERENCES

Ali, A.M., Z. Melegy, M. Morsy, R.M. Megahid, T. Bucherl and E.H. Lehmann, 2004. Image reconstruction techniques using projection data from transmission method. Annals of Nuclear Energy (ELSEVIER), 31: 1415-1428.

Becht, A., J. Tronickez, E. Appel and P. Dietrich, 2004. Inversion strategy in crosshole radar tomography using information of data subsets. Geophysics, 69: 222-230.

Bioshop, T.N., K.P. Bube, R.T. Cutler, R.T. Langan, P.L. Love, J.R. Resnik, R.T. Shuey, D.A. Spindler and H.W. Wyld, 1985. Tomographic determination of velocity and depth in laterally varying media. Geophysics, 50: 903-923.

Bioshop, I. and P. Styles, 1990. Seismic tomographyic imaging of a buried concrete target. Geophysical Prospecting, 38: 169-188.

Chen, S. and S. Billings, 1992. Neural networks for non-linear dynamic systems modelling and identification. Int. J. Control, 56: 319-346.

Elman, J.L., 1990. Finding structure in time. Cognitive Sci., 14: 179-212.

Eric Davalo and Patrick Naïm, 1993. Des réseaux de neurones, Editions Eyrolles.

Groof (de), P.F.M.,1992. Neural networks experiments on synthetic seismic data. Artificial intelligence in the petroleum industry. Edition Technip, pp: 93-124.

Jiang, M. and G. Wang, 2003. Convergence of the Simultaneous Algebraic Reconstruction Technique (SART). IEEE. Trans. Image Proc., 12: 957-961.

Kissling, E., S. Husen and F. Haslinger, 2001. Model parameterization in seismic tomography: A choice of consequence for the solution quality. Physics of the Earth and Planetary Interior (ELSEVIER), 123: 89-101.

Macies, C.C., M.K. Sen and P.L. Stoffa, 2000. Artificial neural network for parameter estimation in geophysics: Geophysical Prospecting, 48: 21-49.

Parisi, R., E.D. Di Claudio, G. Orlandi and B.D. Rao,1996. A generalized learning paradigm exploiting the structure of feedforward neural networks. IEEE. Trans. Neural Networks, 7: 1450-1459.

Peterson, J.E., B.N.P. Paulsson and T.V. McEvilly, 1985. Application of algebraic reconstruction technique to crosshole seismic data. Geophysics, 50: 1566-1580.

Pratt, R.G. and M.H. Worthington, 1988. The application of diffraction tomography to cross-hole seismic data. Geophysics, 53: 1284-1294.

Stewart, R.R., 1991. Exploration seismic tomography: Fundamentals, Course Notes Series, 3: Dominico, S.N. (Ed.), Society of Exploration Geophysicists.