RECOVERY OF PARAMETERS OF A HOMOGENEOUS ELASTIC LAYER USING NEURAL NETWORKS

D. N. Tumakov[*], D. M. Khairullina and A. A. Valeeva

Kazan Federal University, Institute of Computational Mathematics and Information Technologies, 18 Kremlevskaya Str., Kazan 420008, Russia

ABSTRACT

The problem of recovery of parameters of an elastic uniform layer is considered. The solution method of this problem is proposed by means of neural networks. For the input network parameters, amplitudes of transmitted or reflected waves are chosen. We consider the output parameter of network density and longitudinal speed of the environment filling a layer. Three types of neurons activation function are chosen: piecewise-linear, sigmoid and radial function (Gauss's function). The following algorithms of perseptron training are considered: the backpropagation algorithm and the genetic algorithm. In the genetic algorithm, the crossover method, in which the next gene is selected equiprobable and incidental from genes of the ancestors located in the same position is chosen. Then genes of a new chromosome with small probability are exposed to mutations at a small value. Training of neural network takes place according to the calculated data of the direct problem solved multiply times with various elastic parameters varying in the defined intervals. Dependences graphs of an error of parameter recovery on such parameters of neural network as selection dimension and neurons' number are provided. Deterioration in accuracy of recovery of required data, when training network, is shown. Results of comparison of neural networks with various activation functions are described. For density and speed of a layer recovery at the same time, the method of calculation of unknown elastic parameters using the neural network trained for recovery of acoustic rigidity and wave number is proposed. The conclusion is drawn that radial function of neurons activation yields steadier results.

1. INTRODUCTION

Problems of elastic waves passing through isotropic [1] and anisotropic [2, 3] layers arise in many applications. It is often necessary to carry out the nondestructive analysis of the studied layers. In this regard, recovery of above-mentioned parameters of elastic materials according to the available "external" data is of great interest. Most often, such data are amplitudes of the reflected or transmitted waves [4]. The Artificial Neural Network (ANN) can be considered as one of the tools for the similar problems solution [5, 6]. In the current study, for recovery of a profile the method of neural networks is used. For example, authors earlier showed a possibility of the ANN use for recovery of refraction index of a layer [7].

In the current study, ANN training is provided by two methods: by method of the backpropagation and genetic algorithm [8]. At the same time, three activation functions (AF) of neurons are considered: piecewise-linear, sigmoid and radial function (Gauss's function) [9]. Errors of recovery of longitudinal wave speed in a layer for various values of elastic parameters and ANN parameters are estimated. The conclusion is drawn that more exact results are obtained by ANN with radial activation function.

For density and speed of a layer recovery, the method of calculation of unknown elastic parameters by the neural network trained for recovery of acoustic rigidity and wave number is proposed. It is shown that such approach, in certain cases, is more effective than a standard method when the network restores directly through required parameters. Graphs of dependence of neural network training errors on the ANN various parameters are provided.

2. Problem statement

Let the elastic harmonic wave with a frequency of oscillations ω of the form $u_0(x)\exp\{i\omega t\}$, where $u_0(x) = A_0 \exp\{-ik_1 x\}$ and $k_1 = \omega / v_1$ (see Fig. 1) fall on a uniform layer of thickness L (the environment 2 $\{0 < x < L\}$, with unknown density ρ_2 and propagation speed of a longitudinal wave v_2) from the environment 1 $\{x < 0\}$.
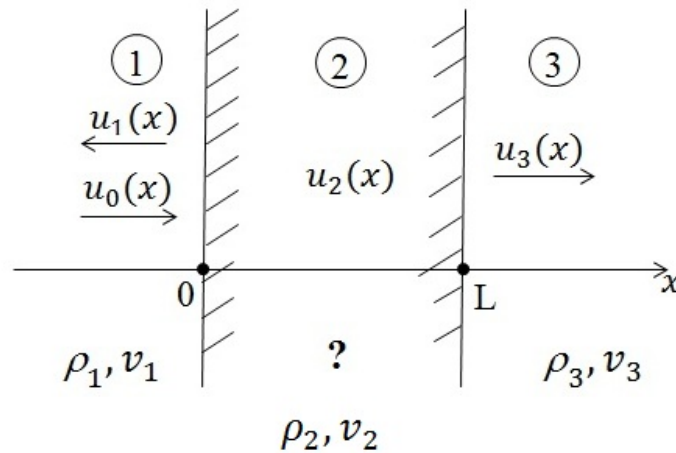
Fig.1. Geometry of the problem

As a result of diffraction, the reflected wave $u_1(x)$ and the wave $u_3(x)$, which passes into the environment 3 $\{x > L\}$ appear. It is necessary to find values of elastic parameters ρ_2 and v_2 provided that the passed and/or reflected waves are known.

Let us solve a problem by the method of neural networks. At the same time, we will consider two methods of network training: the backpropagation algorithm and the genetic algorithm using three most widespread AF: piecewise-linear, sigmoid and radial (Gauss's function).

3. Neural networks

Before proceeding directly to the problem's solution, we will provide some necessary background from the theory of ANN, which we will use further on.

Formally, every neuron represents a certain node, which receives a signal from the previous layer and transfers it to the following layer. The function calculating an output signal of an artificial neuron is called AF. As an argument, the function receives a signal produced at the exit of the so-called adder, which represents the sum of products of weights and values of entrance neurons. It is possible to consider weight as certain channels characterizing a constraint force between neurons and an imitating work of biological synapses. Thus, ANN can be presented in the form of sets of simple elements, i.e. neurons, which are carrying out processing of entrance data with the subsequent delivery of the resulting value. The block diagram of ANN with three hidden layers is presented in Fig. 2.

Most of such functions of activation as piecewise-linear, sigmoid and radial functions and, in addition, some modified options of these functions were widely adopted. A correctly picked AF together with a training algorithm in many respects improve convergence of a neural network. In this regard, one of the important aspects of the ANN design for the solution of a specific objective is the correct choice of the network architecture.
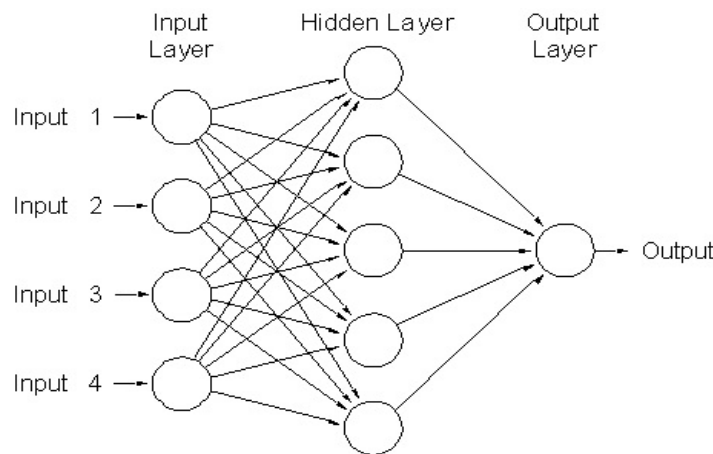
Fig.2. Model of artificial neural network with one hidden layer

The signal at the exit neuron i with the linear activation function is determined by the combined entrance, which represents an adder exit:

$$S_i = \sum_{j=0}^{N} w_{ij} x_j, \tag{1}$$

where w_{ij} are weights of neural network, x_j is the exit of neuron i of the current layer.

Here w_{i0} is value of a threshold, $x_0 = 1$.

Piecewise-linear functions are used most often:

$$f(S) = \begin{cases} 1, S > 1/\alpha, \\ \alpha S, 0 \leq S \leq 1/\alpha, \\ 0, S < 0. \end{cases} \tag{2}$$

Piecewise-linear functions are some of the simplest functions of activation and easily implemented in practice. However, these functions have an essential shortcoming; their derivative has a discontinuity, and it is difficult to use gradient methods for ANN training of this kind.

The sigmoid activation function changes in the range from zero to unity and is described by the following formula:

$$f(S) = \frac{1}{1 + \exp\{-\alpha S\}} \ .$$

(3)

A distinctive property of such function is its monotoneness and differentiability in the entire range of definition that can be used in gradient training algorithms.

It is possible to designate networks of radial and basis functions as a separate class ANN, where radial basis functions are used as activation functions. Gauss's function can act as such function

$$f(S) = \exp\left\{-\frac{S^2}{\sigma^2}\right\},$$

(4)

where parameter σ is a real number.

While determining a structure of the neural network, it is necessary to find values of weights w_{ij} and thresholds w_{i0} so that the error at the exit of the neural network is minimized. For the solution of this problem, training algorithms are used. Let us consider two algorithms, which will be used in the present work.

The most widespread method of training is the backpropagation algorithm. The method of calculating a gradient vector for the subsequent process of weights specification is the cornerstone of this algorithm. The direction of a vector corresponds to the shortest descent on the error surface. This method means training with "a teacher" that is the training selection, at which the network will study having entrance data and a desirable response to it. By comparison of real and target (desirable) values, the mean square error of network training ε is calculated. On the basis of ε, the overall performance of ANN, in general, is estimated.

Let us consider other method of global optimization - a genetic algorithm. The genetic algorithm represents the method reflecting natural evolution of the required decision. In it the evolutionary principle of survival of the most adapted individuals is used. Optimization takes place by crossover and mutation of chromosomes $W = \{w_{ij}\}$ of this population. Under a chromosome W it is necessary to understand a set of genes w_{ij}, which can be coded in the form of a vector ("genotype"). It is presumed that the genotype has the fixed length. On the first iteration in a random way, the set of genotypes of initial population is created. The error of network training helps to estimate the level of fitness of each individual, and on the basis of it to create new population of individuals. The best individuals on the basis of fitness value to

whom operations of crossover and a mutation are applied again get out of the received set of the decisions called by generations. Descendants have to have an opportunity to inherit signs of both parents.

"Mixing" can be carried out in various ways. In the current study we will consider a way of crossover at which the next gene gets out equiprobably and incidentally of genes of the ancestors located in the same position. Further genes of the "designed" chromosome with small probability are exposed to mutations at a small rate. Generation of new populations happens until fitness function is optimized. Thus, the sequence of similar iterations can simulate "evolutionary process".

4. Recovery of longitudinal wave speed in a layer

Let us consider a case when density ρ_2 and thickness of a layer L is known. Let us assume that measurements on both sides of a layer are performed. Results of measurements can be presented as two complex values: $u_1(0)$ and $u_3(L)$. Thus, the designed neural network will contain four neurons at an entrance $\text{Re}[u_1(0)]$, $\text{Im}[u_1(0)]$, $\text{Re}[u_3(L)]$, $\text{Im}[u_3(L)]$, and one neuron at the exit (v_2).

We will train ANN selection containing M experiments. Each experiment corresponds to the value $v_2^{(i)}$, $i = 1..M$, so that the interval $[v_2^{(1)}, v_2^{(M)}]$ contained the restored required value v_2. And we will cover an interval with a uniform grid:

$$v_2^{(i)} = (i-1)\frac{v_2^{(M)} - v_2^{(1)}}{M-1} + v_2^{(1)}.$$ The research of the constructed ANN shows that function of error has many local extrema, and backpropagation algorithm will stop in the next local minimum. In this regard genetic algorithm should be taken into account as a method of global optimization.

At each ANN training as a result we will receive the set full set of weights $W = \{w_{ij}\}$. And owing to any starting values $W^{(0)}$, as well as variabilities of mutations of weights W at each training constructed weight W will differ. Let us assume that the values v_2 calculated by neural network approximately satisfy to the normal law. Then it will allow using the following approach. For the same experiments we will provide training of K times in

network, having received, thus, actually K independent ANN. Further we will calculate exits for each network.

We will receive total value as an average

$$v_2 = \frac{1}{K} \sum_{i=1}^{K} v_{2,i},$$ (5)

For numerical experiments we will choose $K = 5$.

For the ANN "work" accuracy check we use a method of cross validation. We use the following modification of a method. We have available M experiments. Randomly we will choose from them one experiment, we will train network at the remained $M - 1$ data. Let us compare v_2 received with the known value, having received an approach error ε_i. Let us carry out R tests, we will be result the sequence $\varepsilon_i, ..., \varepsilon_R$. Let us choose the worst value $\varepsilon = \max_{i=1..R} \varepsilon_i$. Let $R = 20$, then it is possible to claim 95% with reliability that the neural network "works" with a margin ε error. Let us put $R = 20$.
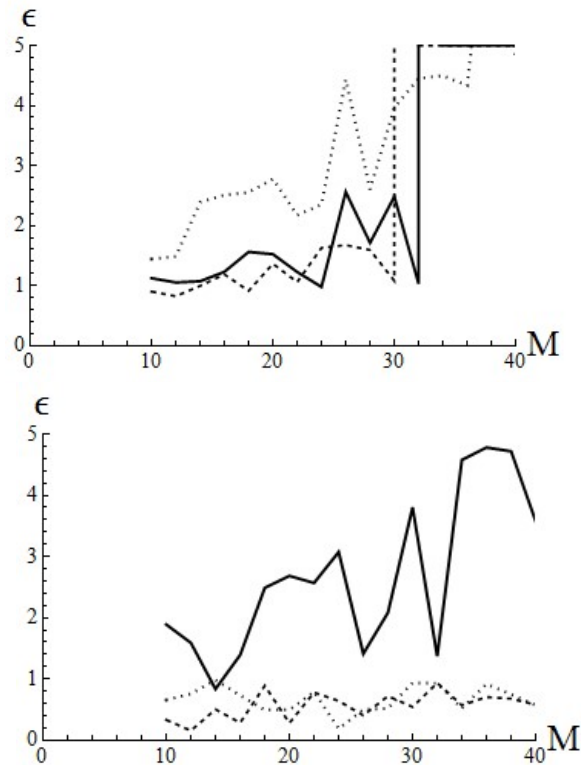
Fig.3. Dependence of ε on M. ANN: 1 hidden layer with 3 neurons (left), 3 hidden layers with 3 neurons (right). The continuous line is for the piecewise-linear activation function, the dashed line is for sigmoid function, the dotted line is for Gauss's function. The backpropagation algorithm was applied

Let us consider recovery of speed on transmitted field. In this case the neural network will contain two entrances $\mathrm{Re}[u_3(L)]$, $\mathrm{Im}[u_3(L)]$ and one exit v_2. Let us restore value v_2 the network containing one and three hidden layers on three neurons in each. Dependence of an error ε on number of selections M is given in Fig. 3. It is visible that the error remains rather big for ANN with one hidden layer with three neurons, but steadier and a little smaller values ε will be for ANN with three hidden layers with three neurons in each layer. It should be noted that at big M there is an instability of an error caused by retraining.
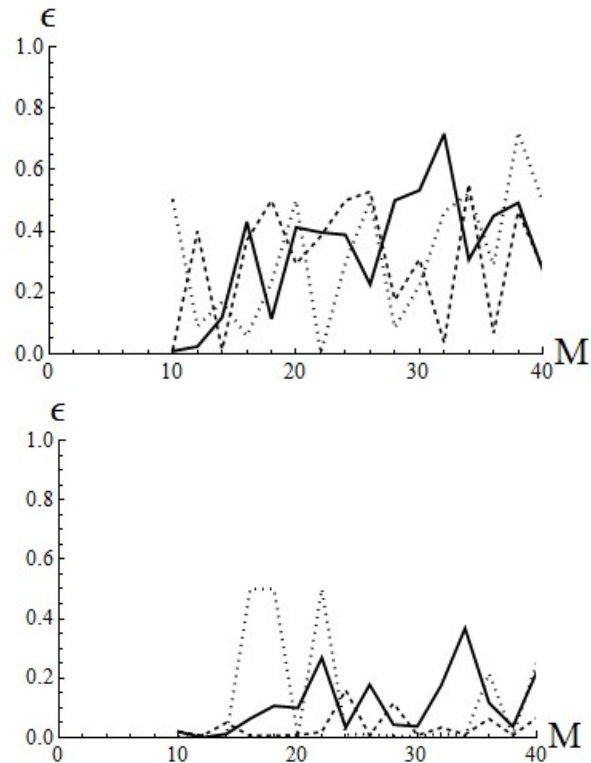
Fig.4. Dependence of ε on M. ANN: 1 hidden layer with 3 neurons (left), 3 hidden layers with 3 neurons (right). The continuous line is for the piecewise-linear activation function, the dashed line is for sigmoid function, the dotted line is for Gauss's function. Genetic algorithm was applied

We will make the next experiments for a genetic algorithm (Fig. 4). It is obvious that the error improves and stabilizes with growth of quantity of layers and neurons (and according to weights). In these cases it is already difficult to allocate any activation function.
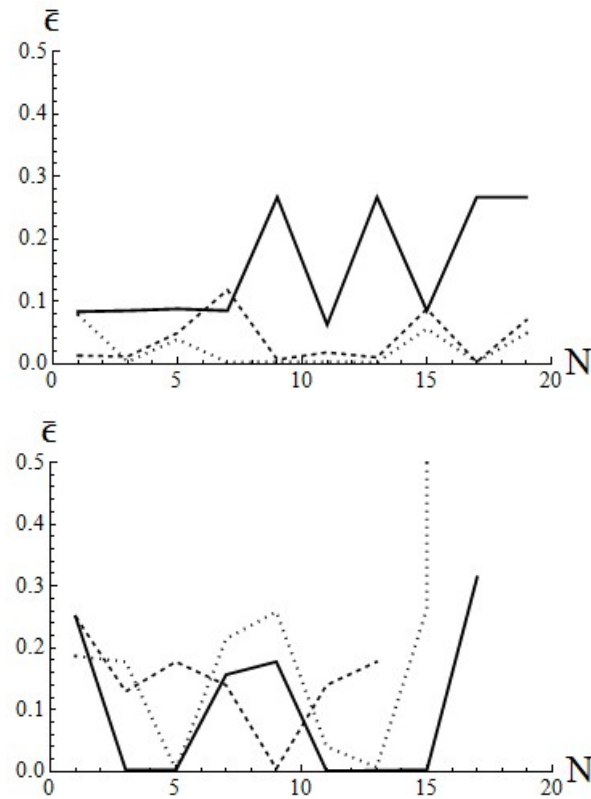
Fig.5. Dependence $\bar{\varepsilon}$ on N. Activation function of output neuron of the transmitted layer: identical (left), linear (right). The continuous line is for the piecewise-linear activation function, the dashed line is for sigmoid function, the dotted line is for Gauss's function. The backpropagation algorithm was applied

In Fig. 5 dependence of a relative error recovery of longitudinal wave speed is given in a layer $\bar{\varepsilon}$ from number of neurons N of network with one hidden layer. It is possible to see from graphs, when using the AF gradient methods of output neuron of the transmitted layer to set better identical with AF of neurons of inside layers. In case of a linear exit results turn out unstable.

Let us consider the same dependences for a genetic algorithm of network training (see Fig. 6). It is visible that $N > 7$ at values of error $\bar{\varepsilon}$ with growth N improves and stabilized, at the same time both options have approximately identical results. We will receive the best result when using AF of Gauss with a linear exit.
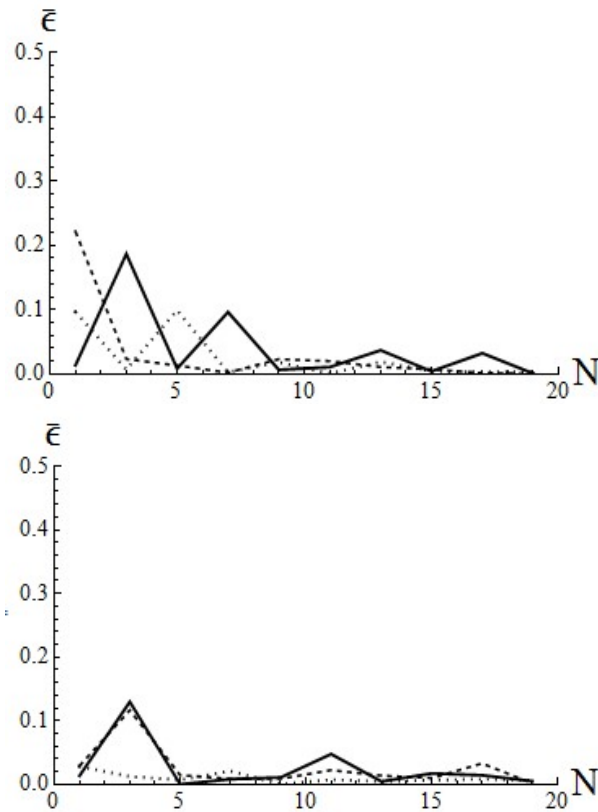
Fig.6. Dependence $\bar{\varepsilon}$ from N. Activation function of output neuron of the transmitted layer: identical (left), linear (right). The continuous line is for the piecewise-linear activation function, the dashed line is for sigmoid function, the dotted line is for Gauss's function. Genetic algorithm was applied

By method of cross validation, it is expedient to reveal dependence of the training error on amount of neurons of the network with one hidden layer. Conditions for cross validation remain same, as for assessment of the training error dependence on number of selections M. These graphs for the backpropagation algorithm and a genetic algorithm are provided on Fig. 7.
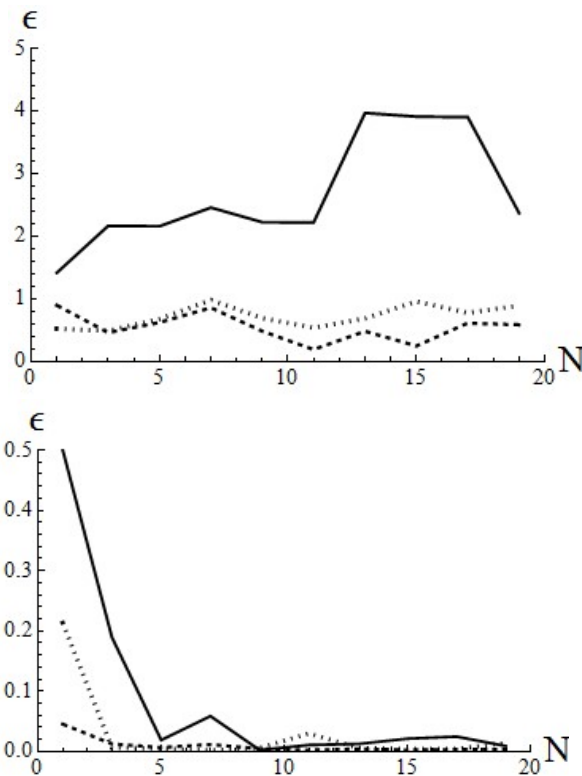
Fig.7. Dependence ε from N. The backpropagation algorithm (left), the genetic algorithm (right). The continuous line is for the piecewise-linear activation function, the dashed line is for sigmoid function, the dotted line is for Gauss's function. Activation function of output neuron of the transmitted layer identical

For a genetic algorithm with sigmoid function and Gauss's function the error of network training, starting $N = 3$, decreases and further essential jumps are not observed. If to look at error change ranges, by the backpropagation algorithm the network works with a bigger error, especially for piecewise-linear activation function.

5. Recovery of longitudinal speed and density

Let us consider more general problem - the problem of the two parameters recovery of an elastic layer: longitudinal speed v_2 and density ρ_2. At the same time we will consider two approaches depending on a way of parameter recovery. In the first case the neural network restores values v_2 and ρ_2. In the second restores rigidity of the environment $z_2 = v_2\rho_2$ and the wave number $k_2 = 2\pi\omega/v_2$, and required speed and density are expressed through them.

We choose the number of the training selection $M = Mv * M\rho$, where Mv and $M\rho$ are the grid values on v_2 and ρ_2, respectively. Let us note that in the second case the grid on k_2 and z_2 turns out curvilinear as these parameters are not linearly connected with v_2 and ρ_2.
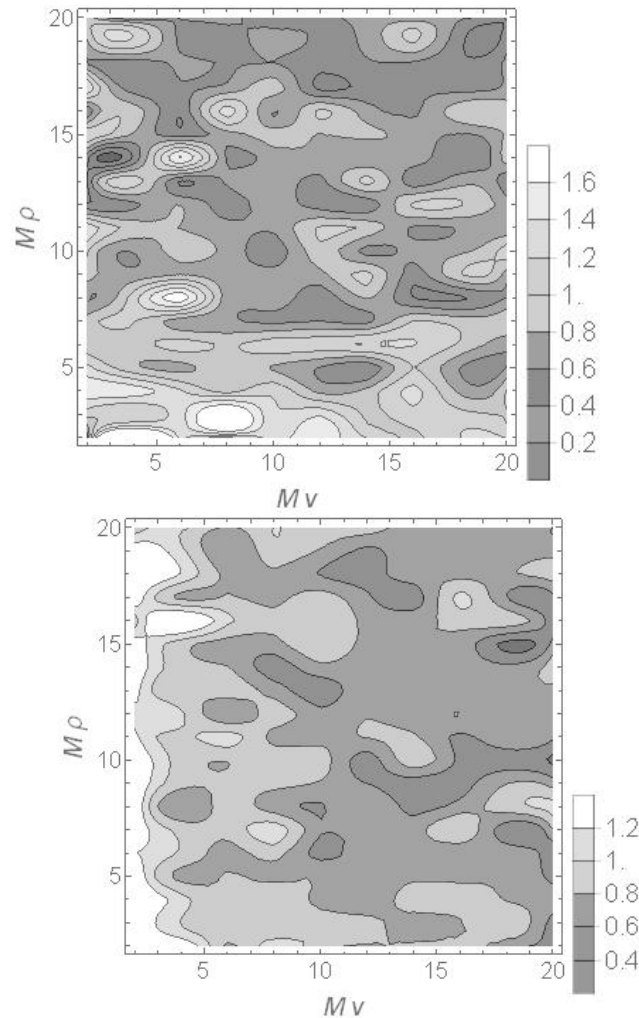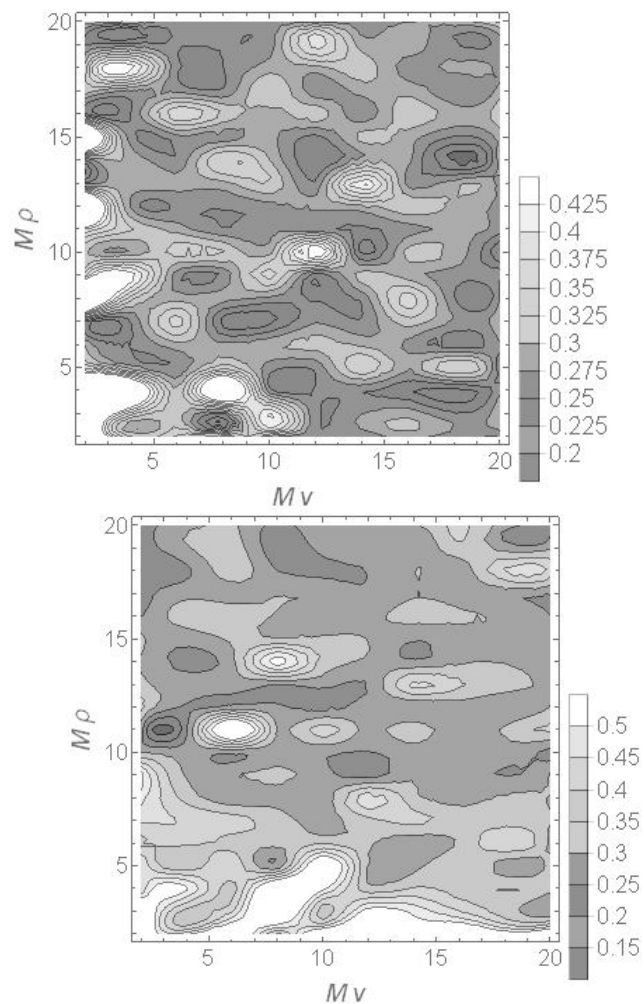

Fig.8. Relative error of recovery of v_2. Error at recovery: through z_2, k_2 (left), v_2, ρ_2 (right). The sigmoid AF from the linear AF in an output layer. Genetic algorithm was applied

Let us make numerical experiments for a case $K = 3$. Lines of level of an error recovery for longitudinal speed are given in two ways in Fig. 8: through v_2, ρ_2 (right) and z_2, k_2 (left). The relative mistake in both cases remains rather big - more than 100%. Nevertheless, recovery through v_2, ρ_2 yields better results, than recovery through k_2 and z_2. It is visible in the right Fig. 8 that the worst approach $\bar{\varepsilon} > 1.2$ turns out at small values Mv.

While at the second way of recovery it is impossible to allocate specific zones of bad speed recovery.

Now we will consider dependence of a relative error recovery ρ_2 on selection M number for ANN from sigmoid AF in inside layers and linear AF in an output layer. Lines of level of an error recovery are also given in two ways in Fig. 9: through v_2, ρ_2 (right) and z_2, k_2 (left). Also as well as at recovery of speed, big mistakes are located at small values $M\rho$. Results of experiments of recovery ρ_2 (see Fig. 9) demonstrate that the second way of recovery is a little more preferable.

Fig.9. Relative error of recovery of ρ_2. Error at recovery: through z_2, k_2 (left), v_2, ρ_2 (right). The sigmoid AF from the linear AF in an output layer. Genetic algorithm was applied

For more detailed comparison of two ways of recovery we will estimate a difference of errors for each of them. For this purpose we will estimate a difference between error at recovery in the second way $\bar{\varepsilon}(z_2, k_2)$ and error at recovery in the first way $\bar{\varepsilon}(v_2, \rho_2)$. Results of values $\bar{\varepsilon}(z_2, k_2) - \bar{\varepsilon}(v_2, \rho_2)$ for ANN with two various AF are given in Fig. 9 and 10. Zones with negative values show where the second way of recovery of required parameters is more preferable. While zones with positive values indicate higher precision at recovery in the first way. Parameter recovery in all cases was carried out by ANN with two hidden layers on three neurons in each.
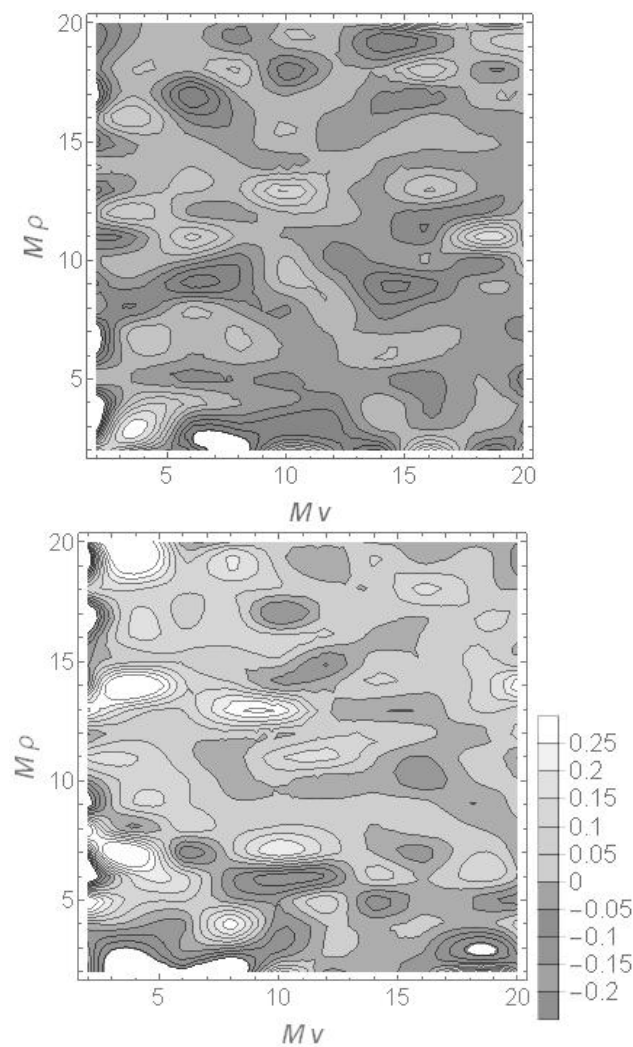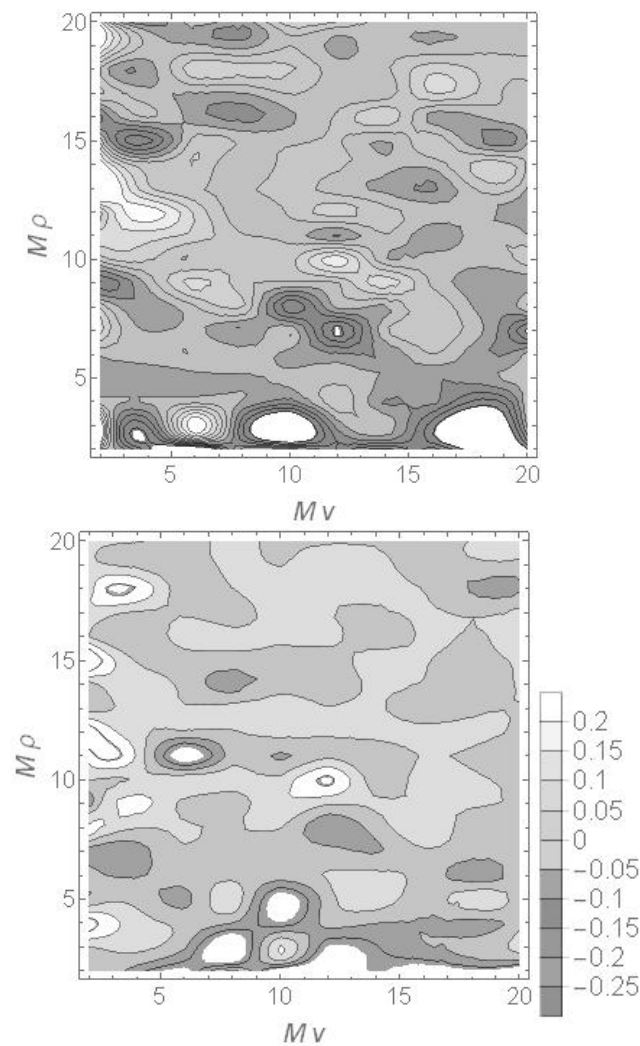

Fig.10. A difference of results at recovery ρ_2 as $\bar{\varepsilon}(z_2, k_2) - \bar{\varepsilon}(v_2, \rho_2)$ for ANN with Gauss's AF from identical AF of an output layer (left) and from the linear AF (right). Genetic algorithm was applied

The difference of error is given in Fig. 10 at various Mv and $M\rho$ for neural network with Gauss's AF. Left lines of level for a case when all layers have one AF are given; and right - for a case from Gaussian AF in the hidden layers and linear AF in an output layer.

It is possible to note prevalence of dark color in the left Fig. 10. It demonstrates that ANN from Gaussian AF restores better on z_2 and k_2. Dark and light shades in the right Fig. 10 are distributed approximately equally. It means that both ways for neural networks from linear AF in an output layer give an identical error.

Fig.11. A difference of results at recovery ρ_2 as $\bar{\varepsilon}(z_2, k_2) - \bar{\varepsilon}(v_2, \rho_2)$ for ANN from sigmoid AF from identical AF of an output layer (left) and from the linear AF (right). Genetic algorithm was applied

For sigmoid AF dark shades in Fig. 11 are expressed not so strongly. This fact shows that in this case it is impossible to give preference to any way.

Thus, if to estimate an error of speed recovery, then small preference should be given to the first way - on density and speed. If to estimate an error of density recovery, then the second way recovery is more preferable to radial AF through z_2 and k_2.

6. CONCLUSIONS

It is expedient to provide training of the neural network for recovery of elastic parameters of wave propagation in a uniform layer by means of the genetic algorithm in case of a simple ANN. Such conclusion follows from the fact that the criterion function (error of the network) contains a set of local extrema.

The second and quite obvious conclusion is that an increase in the number of neurons improves an approximation of the required values, and the network functioning with a complication of the structure (an increase in the number of neurons) becomes more stable.

The third conclusion concerning functions of activation demonstrates that all functions considered in the present work (piecewise-linear, sigmoid and Gauss's functions) approximate the sought-for speeds almost identically. But an ANN giving more stable solutions is obtained via using a genetic algorithm with the Gaussian activation function.

For recovery of two parameters, which are longitudinal speed and density, it is more preferable to use the network trained for recovery of acoustic rigidity and wave number.

Let us note that in the present study, we were not concerned with questions of uniqueness of the solution of the inverse problem and optimum frequencies for the analysis. However, we fully understand that this question is quite important [10, 11] and deserves a separate attention and research investigation. In addition, it is better to train the network via taking into account the features of passing of elastic waves through characteristic peculiarities of the waveguide structures [12, 13].

7. SUMMARY

Perseptrones with a small number of neurons and a high precision recover one unknown parameter of an elastic layer and can be recommended as a method of fast and qualitative analysis.

Recovery of two unknown parameters such as density and speeds of propagation of an elastic wave provide a solution to the problem with a great error. This approach can be used as an approximate or initial assessment of the elastic parameters.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. V. Anufrieva, D. N. Tumakov and V. L. Kipot, "Elastic wave propagation through a layer with graded-index distribution of density," in *Proc. DD'2012*, pp.21-26, 2012.

[2] A. V. Anufrieva and D. N. Tumakov, "Diffraction of a plane elastic wave by a gradient transversely isotropic layer," *Advances in Acoustics and Vibration*, vol. 2013, 262067, 2013.

[3] A. Anufrieva, D. Chikrin and D. Tumakov, "On Peculiarities of Propagation of a Plane Elastic Wave through a Gradient Anisotropic Layer," *Advances in Acoustics and Vibration*, vol. 2015, 515263, 2015

[4] N. B. Pleshchinskii and D. N. Tumakov, "The reconstruction of dielectric profile of a layer for the harmonic wave case," in *Proc. PIERS 2013*, pp.643-647, 2013.

[5] G. Caviglia and A. Morro, "Inversion of reflection data in an isotropic multilayered medium," *Acta Mechanica,* 189(1–4), pp. 65-72, 2007.

[6] V. Brovko, E. K. Murphy and V. V. Yakovlev, "A neural network technique for reconstruction of 2D complex permittivity profiles of materials in waveguide systems," in *Proceedings of the Global Congresson Microwave Energy Applications, Otsu, Japan,* pp.381-384, 2008.

[7] D. N. Tumakov and D. M. Khairullina, "Application of neural network method to restore the refraction index of homogeneous dielectric layer," *Research Journal of Applied Sciences*, vol. 10, no. 8, pp. 419-427, 2015.

[8] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks,* vol. 61, pp. 85-117, 2015.

[9] S. Haykin *Neural Networks and Learning Machines* / S. Haykin. −Prentice Hall, 2009.

[10] D. Tumakov, "On uniqueness of a solution to the problem of reconstruction of the homogeneous dielectric layer parameters," *Far East Journal of Mathematical Sciences*, vol. 101, no. 4, pp. 763-772, 2017.

[11] D. Tumakov, "On optimal frequencies for reconstruction of a one-dimensional profile of gradient layer's refractive index," *International Journal of Optics,* vol. 2014, 841960, 2014.

[12] A. V. Anufrieva, D. N. Tumakov and V. L. Kipot, "Peculiarities of propagation of a plane elastic wave through a gradient layer," in *Proc. DD'2013*, pp. 11–16, 2013.

[13] A. V. Anufrieva and D. N. Tumakov, "On some of the peculiarities of propagation of an elastic wave through a gradient transversely isotropic layer," in *Proc. DD'2014*, pp. 23–28, 2014.