

PARAMETERS INVERSION OF FLUID-SATURATED POROUS MEDIA BASED ON NEURAL NETWORKS^{*}

Wei Peijun^{1,2} Zhang Zimao¹ Han Hua¹

(¹ *Institute of Mechanics, Northern Jiaotong University, Beijing 100044, China*)

(² *LNM, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100080, China*)

ABSTRACT The multi-layers feedforward neural network is used for inversion of material constants of fluid-saturated porous media. The direct analysis of fluid-saturated porous media is carried out with the boundary element method. The dynamic displacement responses obtained from direct analysis for prescribed material parameters constitute the sample sets training neural network. By virtue of the effective L-M training algorithm and the Tikhonov regularization method as well as the GCV method for an appropriate selection of regularization parameter, the inverse mapping from dynamic displacement responses to material constants is performed. Numerical examples demonstrate the validity of the neural network method.

KEY WORDS fluid-saturated porous media, parameter inversion, neural networks, boundary elements

I . INTRODUCTION

The inverse problem of material constants is of interest in rock and soil engineering and geophysical prospecting engineering. The approximations of material parameters are generally obtained by inverse analysis of the static and dynamic displacement response measured in situ. The classical inverse methods reduce an inverse problem of material parameters to an optimal problem by fitting displacement simulated with the displacement measured^[1]. The solution of an optimal problem is obtained generally by searching in the space of parameters. Deterministic searching and stochastic searching are two main searching methods. The information of derivative of one-order or two-order is needed in deterministic searching methods, such as the gradient method and Newton method. Unfortunately, the extraction of gradient information is often difficult and time-consuming, especially for multi-parameters inversion problems. The computation of gradient by the time-convolution regularization method^[2] or the time-inverse wave field method^[3] shows the intricacy. Furthermore, the derivative-based searching methods will fall in the local extremum of misfit error function easily. The stochastic searching method, for example, simulated anneal algorithms^[4] and genetic algorithms^[5], need only the information of misfit error function by adoption of Metropolis' criterion and criterion of survival of the fittest. Therefore, the time-consuming computation of gradients is avoided. But the effectiveness of the stochastic searching method is sensitive to the selection of various probability parameters. The neural network method is a powerful tool for dealing with complicated nonlinear problems. The feedforward networks of 3 layers with sufficient neural elements can perform any complicated nonlinear mapping. Stavroulakis and Antes^[6] used the neural network approach to deal with crack detection problems recently. It stands to rea-

* Project supported by the National Natural Science Foundation of China (Nos. 19872002 and 10272003) and Climbing Foundation of Northern Jiaotong University.

Received 10 December 2001; revision received 29 September 2002.

son that neural networks can perform mapping displacements response measured to material constants and the key problem is how to design the structure of neural networks and how to train neural networks. It's feasible to design the structure of neural networks with genetic algorithms, although great cost has to be paid^[7]. But it's more common to design the structure of neural networks by trial-and-error and using the experience of the designer^[8]. The methods of training neural networks have been studied widely after the back propagation algorithm is proposed, for example, the variable learning rate method, conjugate gradient algorithm and quasi-Newton algorithm. The L-M algorithm appears to be the fastest method among these methods because it's designed to approach second-order training speed without having to compute the Hessian matrix^[9]. The sample sets training network can be obtained from direct computation by using various numerical methods, for example, finite element methods and boundary element methods.

The parameters inverse problem of fluid-saturated porous media is studied by using the neural network approach in this paper. For the semi-space of media often encountered in geophysical prospecting problems, the finite element method brings inevitably the introduction of complicated artificial boundary. Alternatively, the infinite boundary elements are introduced when the boundary element method is used. In contrast, the boundary element method is more suitable for semi-space problems owing to the relative simplicity of infinite boundary element. The basic solutions of fluid-saturated porous media are necessary for boundary element methods and can be obtained with the help of analogy between poroelasticity and thermoelasticity^[10]. A feedforward neural network of 3 layers is used to perform the mapping from displacement response to material constants. The sample data sets for training neural networks are obtained from direct computation by the boundary element method. And the effective L-M method is used to train neural networks. Two numerical examples are given to show the validity of the neural network method.

II . THE BOUNDARY ELEMENT METHOD FOR FLUID-SATURATED POROUS MEDIA

The Biot's equations of fluid-saturated porous media are expressed as

$$t_{ij,j} + X_i = \rho_{11} \ddot{u}_i + \rho_{12} \ddot{U}_i + b(\dot{u}_i - \dot{U}_i) \quad (1)$$

$$\tau_{,i} + X'_i = \rho_{12} \ddot{u}_i + \rho_{22} \ddot{U}_i - b(\dot{u}_i - \dot{U}_i) \quad (2)$$

where $t_{ij} = (\lambda + Q^2/R) \delta_{ij} e + 2\mu e_{ij} + Q \delta_{ij} \theta$, $\tau = Qe + R\epsilon$, $\rho_{11} = (1 - \beta) \rho_s + \rho_a$, $\rho_{22} = \beta \rho_f + \rho_a$, $\rho_{12} = -\rho_a$, ρ_s and ρ_f are mass density of solid phase and fluid phase respectively, ρ_a additional mass density which represents a mass coupling between fluid and solid, β porosity, λ , μ , Q and R material constants, b a dissipation constant, u_i and U_i solid and fluid displacement respectively, e and θ solid and fluid dilation, X_i and X'_i body forces acting on solid and fluid.

From Betti's reciprocal relation, the following equation holds

$$\begin{aligned} & \int_{\Gamma} (t_i u_i^* + \tau U_n^*) d\Gamma + \int_{\Omega} (X_i u_i^* + X'_i U_i^*) d\Omega \\ &= \int_{\Gamma} (t_i^* u_i + \tau^* U_n) d\Gamma + \int_{\Omega} (X_i^* u_i + X_i'^* U_i) d\Omega \end{aligned} \quad (3)$$

where Γ is the boundary of domain Ω , $t_i = t_{ij} n_j$ and $U_n = U_i n_i$, \mathbf{n} is the unit normal to the boundary.

For dynamic poroelastic problem in frequent domain, by using the fundamental solutions corresponding to the following two sets of body forces^[10]

$$(1) X_i^* = \delta(\mathbf{x} - \boldsymbol{\xi}) \delta_{ij}; X_i'^* = 0$$

$$(2) X_i'^* = [(1/2\pi) \ln r]_{,i}; X_i^* = X_i'^* (i\omega b + \omega^2 \rho_{12}) / (-i\omega b + \omega^2 \rho_{22})$$

The integral representations of the solid displacement u_i and fluid pressure τ are obtained in case of body force being ignored^[10]

$$c_{ij} u_i + \int_{\Gamma} t_{ij}^* u_i d\Gamma + \int_{\Gamma} \tau_j^* U_n d\Gamma = \int_{\Gamma} (t_{ij} u_{ij}^* + \tau U_{nj}^*) d\Gamma \quad (i, j = 1, 2) \quad (4a)$$

$$Jc_{33} \tau + \int_{\Gamma} u_{i3}^* t_i d\Gamma + \int_{\Gamma} \tau (U_{n3}^* - JX_i' n_i) d\Gamma = \int_{\Gamma} (t_{i3}^* u_i + \tau_3^* U_n) d\Gamma \quad (4b)$$

$$\text{or} \quad \mathbf{c}^i \mathbf{u}^i + \int_{\Gamma} \mathbf{p}^* \mathbf{u} d\Gamma = \int_{\Gamma} \mathbf{u}^* \mathbf{p} d\Gamma \quad (5)$$

where $\mathbf{u} = \{u_1, u_2, \tau\}^T$, $\mathbf{p} = \{t_1, t_2, U_n\}^T$

$$\mathbf{p}^* = \begin{bmatrix} t_{11}^* & t_{21}^* & -U_{n1}^* \\ t_{12}^* & t_{22}^* & -U_{n2}^* \\ t_{13}^* & t_{23}^* & -U_{n3}^* \end{bmatrix}, \quad \mathbf{u}^* = \begin{bmatrix} u_{11}^* & u_{21}^* & -\tau_{n1}^* \\ u_{12}^* & u_{22}^* & -\tau_{n2}^* \\ u_{13}^* & u_{23}^* & -\tau_{n3}^* \end{bmatrix}, \quad \mathbf{c}^i = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -J \end{bmatrix}$$

Equation (5) can be discretized with boundary elements

$$\begin{aligned} \mathbf{c}^i \mathbf{u}^i + \sum_{l=1}^M \sum_{j=1}^N \left\{ \int_{\Gamma_l} \mathbf{p}^* (\boldsymbol{\xi}, \eta) N_j (\boldsymbol{\xi}, \eta) \mathbf{J} (\boldsymbol{\xi}, \eta) d\Gamma \right\} \mathbf{u}^{(l,j)} \\ = \sum_{l=1}^M \sum_{j=1}^N \left\{ \int_{\Gamma_l} \mathbf{u}^* (\boldsymbol{\xi}, \eta) N_j (\boldsymbol{\xi}, \eta) \mathbf{J} (\boldsymbol{\xi}, \eta) d\Gamma \right\} \mathbf{p}^{(l,j)} \end{aligned} \quad (6)$$

$$\text{or} \quad \mathbf{c}^i \mathbf{u}^i + \sum_{l=1}^M \sum_{j=1}^N \hat{\mathbf{H}}^{ij} \mathbf{u}^{(l,j)} = \sum_{l=1}^M \sum_{j=1}^N \mathbf{G}^{ij} \mathbf{p}^{(l,j)} \quad (7)$$

where M is the number of boundary elements, N the node number on boundary, $N_j(\boldsymbol{\xi}, \eta)$ the shape function, and $\mathbf{J}(\boldsymbol{\xi}, \eta)$ the Jacobian matrix.

Let $\mathbf{H}^{ij} = \hat{\mathbf{H}}^{ij} + \mathbf{c}^i$, and Eq. (7) can be written in a more compact form:

$$\mathbf{H} \mathbf{u} = \mathbf{G} \mathbf{p} \quad (8)$$

Equation (8) can be reordered as usual passing all the unknowns to the left-hand side and the known to the right, then the linear algebraic equations are obtained

$$\mathbf{A} \mathbf{X} = \mathbf{F} \quad (9)$$

where \mathbf{X} is the vector of the unknown boundary value of u and τ , and \mathbf{F} is obtained by multiplying the corresponding columns of \mathbf{H} or \mathbf{G} by the known values of u and τ .

III. NEURAL NETWORK FOR INVERSE MAPPING

By using the boundary element method discussed above, the dynamic displacement response at the surface of fluid-saturated porous media whose material constants are prescribed can be estimated. The process is called direct analysis or direct mapping. The process to inverse determine the material constants from known dynamic displacement response is called inverse analysis or inverse mapping. It's expected that the inverse mapping can be simulated by multi-layers feedforward neural network in order to avoid directly solving the mathematical model of inverse system that is very complicated for

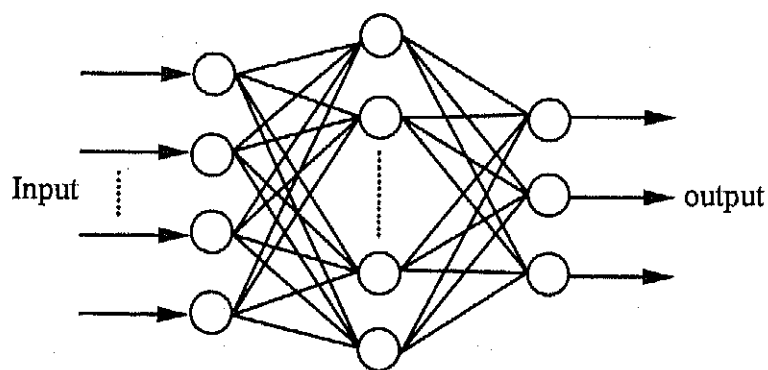


Fig. 1 Inverse mapping neural network for parameter inversion of fluid-saturated porous media.

fluid-saturated porous media. The multi-layers feedforward neural network is composed of a large number of highly interconnected neural elements which are assigned to a number of consecutive layers, as shown in Fig.1. The first layer receives signal and the last layer outputs response. An internal signal is transmitted from one layer to the next in the hierarchy. Each neural element sums the input from the neural elements of the previous layer to which it is connected, magnifies by an appropriate weight which resembles the synaptic strength in neural connections, then transforms it by a transfer function (activation function or response function), and transmits it to the neural elements of the next layer. The numbers of hidden layers and neural elements in the hidden layers are related to the problem studied. The transfer function of neuron in the hidden layer is continuous Sigmoid function.

$$f(x) = 1 / \{1 + \exp[-(x - \theta)]\} \quad (10)$$

where θ is the valve value of neuron. The transfer function of neuron in the output layer is taken as a linear function. The mapping character of the network is determined by the weight W_{ij} between neurons of dissimilar layers and the valve value θ_i of neurons. The net inputs and the outputs of neuron i in layer $k + 1$ in the neural network of M layers can be expressed as

$$n_i^{k+1} = \sum_{j=1}^{S_k} w_{ij}^{k+1} o_j^k + \theta_i^{k+1} \quad (11)$$

$$o_i^{k+1} = f^{k+1}(n_i^{k+1}) \quad (k = 0, 1, \dots, M - 1) \quad (12)$$

Training network is to adjusted gradually weight W_{ij} and valve value θ_i to minimize error function of neural networks by appropriately constructed algorithms according to a large number of data diploid $\{(p_i, t_i), (i = 1, \dots, Q)\}$ obtained from direct analysis, where displacement response p_i correspond to prescribed material constants t_i . Usually, the least mean square error is taken as the error function or misfit error function.

$$E = \sum_{q=1}^Q (t_q - o_q^M)^T (t_q - o_q^M) = \sum_{q=1}^Q e_q^T e_q \quad (13)$$

Among various algorithms of training networks, the L-M algorithm appears to be the faster algorithm for training moderate-sized feedforward neural networks (up to several hundred of weights) because it can be adjusted between gradient method and Newton method by an appropriate factor μ . The network parameters x , including weights W_{ij} and valve value θ_i , is adjusted as following in the L-M algorithm.

$$\Delta x = [J^T(x)J(x) + \mu I]^{-1} J^T(x)e(x) \quad (14)$$

where $J(x)$ is Jacobian matrix composed of one-order derivatives. By considering

$$\nabla E(x) = J^T(x)e(x), \quad \nabla^2 E(x) \approx J^T(x)J(x) \quad (15)$$

it is obvious that the L-M algorithm approaches gradient algorithm when parameter μ is large enough and Gauss-Newton algorithm when parameter μ is small enough. In the process of training, the parameter μ is adjusted as following: if $E[x(k+1)] \geq E[x(k)]$, then $\mu = 10\mu$; else $\mu = \mu/10$.

In order to overcome overfitting which may degrades the prediction of network, the misfit error function is regularized to make the output of network smoother.

$$\hat{E} = \sum_{q=1}^Q e_q^T(x)e_q(x) + \gamma \sum_{j=1}^N x_j^2 \quad (16)$$

where γ is Tikhonov regularization parameter. In case that regularization parameter γ is too small, the purpose of smoothening output of network could not be carried out; in case that regularization parameter γ is too large, the total misfit error would increases apparently. The appropriate regularization parameter γ is given by using general cross validation (GCV) method^[11] to achieve the best tradeoff between smoothening output of network and decreasing misfit error.

IV. NUMERICAL EXAMPLE

The first numerical example studied in the present paper deals with a geomechanical problem of semi-space of fluid-saturated porous media excited by a uniform harmonic normal traction on the surface shown in Fig.2. The amplitude, breadth and frequency of the uniform harmonic normal traction applied on solid phase are: $P_0 = 100$ KPa, $2a = 2$ m and $\omega = 1$ rad/sec. The surface of fluid-saturated porous media is drained. The boundary conditions are expressed as

$$\begin{aligned} \tau_z &= 0, & x &\notin (-a, a) \\ \tau_z &= P_0 e^{i\omega t}, & x &\in (-a, a), \quad \text{at } z = 0 \\ \tau_{zx} &= 0, & x &\in (-\infty, \infty), \quad \text{at } z = 0 \\ \tau &= 0, & x &\in (-\infty, \infty), \quad \text{at } z = 0 \end{aligned}$$

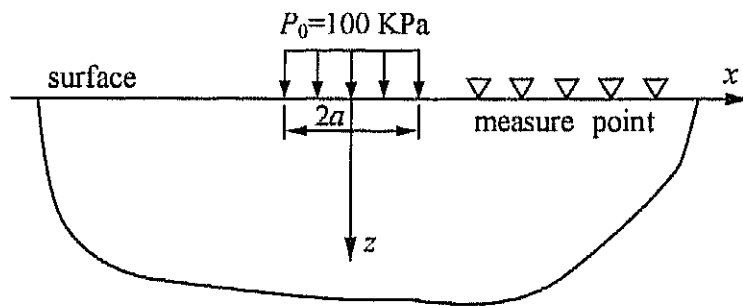


Fig.2 2-D semi-space of fluid-saturated porous media excited by uniform harmonic load.

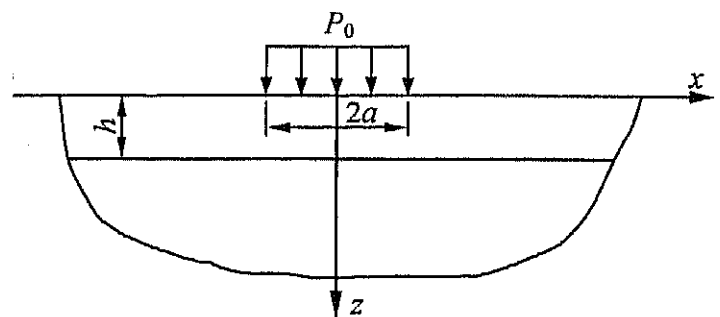


Fig.3 2-D semi-space of fluid-saturated porous media with a fluid-saturated overlay of dissimilar material constants excited by uniform harmonic load.

The reference material constants are: $\lambda = 4 \times 10^9$ N/m², $\mu = 6 \times 10^9$ N/m², $Q = 1.399 \times 10^9$ N/m², $R = 0.444 \times 10^9$ N/m², $b = 0.19 \times 10^9$ Ns/m⁴, $\rho_s = 2700$ kg/m³, $\rho_f = 1000$ kg/m³, $\rho_a = -150$ kg/m³ and $\beta = 0.20$. The unknown material constants are (ρ_s, ρ_f, β) . Fives value are equidistantly collected between 80% and 120% of the reference material constants and 125 sample sets are acquired according the orthogonal design principle. By using boundary element method, the displacement at measured points on surface of fluid-saturated porous media can be computed to obtain the displacement responses to 125 sample sets of material constants. The topological structure of neural network is 10-15-3. 10 input nodes correspond to the horizontal and vertical displacements of 5 measured points and 3 output nodes correspond to 3 unknown material constants. The 125 sample sets are used to train the network with the L-M algorithm. The neural network trained is used to carry out inverse mapping from displacement responses to material constants. In order to check the prediction of network, 5 groups of material constants are obtained randomly and the corresponding displacements are computed by boundary element method. Then, the corresponding displacement responses are put into network. The corresponding material constants are obtained from output of network and are compared with real value in Table 1. The fact that mean errors are not exceed 5% for the 5 groups of material constans indicates that good prediction of material constants can be carried out by network performing inverse mapping through an appropriate tranining of network.

The second numerical example studied is a semi-space of fluid-saturated porous media with a fluid-saturated overlay of dissimilar material constants as shown in Fig.3. The excitation applied on the surface and the boundary conditions are same as in the first example. The reference material constants of the overlay ($h = 5$ m) are: $\lambda = 3 \times 10^9$ N/m², $\mu = 4 \times 10^9$ N/m², $Q = 1.3 \times 10^9$ N/m², $R = 0.4 \times 10^9$ N/m², $b = 0.2 \times 10^9$ Ns/m⁴, $\rho_s = 2100$ kg/m³, $\rho_f = 1000$ kg/m³, $\rho_a = -120$ kg/m³ and $\beta =$

0.4. The unknown material constants are: (1) (ρ_s, ρ_f, β) in the overlay. (2) (ρ_s, ρ_f, β) in the underlying semi-space. The collecting of sample data sets and the training of neural network are same as in the first example. Similarly, 5 groups of material constants obtained randomly and corresponding surface displacements are used to check the prediction of network. The real value of material constants and the prediction of network are list in Table 2 and Table 3, respectively. It is found that prediction of (ρ_s, ρ_f, β) in the overlay is better than that of (ρ_s, ρ_f, β) in the semi-space (the mean relative error is 6.81% for overlay and 9.44% for underlying semi-space). This phenomenon may be explained by that the surface displacement responses are more sensitive to the material constants of overlay than that of fluid-saturated underlying semi-space. In addition, the fact that the mean error in the second example is greater than that in the first example indicates that larger sample data sets are necessary for more complicated problem to reduce the errors.

Table 1 The comparison between be real value and prediction of network for (ρ_s, ρ_f, β) in the semi-space

Test groups	Real value (ρ_s, ρ_f, β)			Predictions of neural network			Relative error (%)		
1	2710	920	0.185	2765	897	0.173	2.03	2.51	6.49
2	3045	834	0.228	2989	876	0.241	1.84	5.04	5.70
3	2846	1176	0.203	2798	1094	0.196	1.69	6.97	3.45
4	3175	1032	0.168	3281	986	0.152	3.34	4.46	9.52
5	2241	856	0.194	2097	795	0.207	6.43	7.13	6.71

* Mean relative error: 4.89(%).

Table 2 The comparison between the real value and prediction of network for (ρ_s, ρ_f, β) in the overlay

Test groups	Real value (ρ_s, ρ_f, β)			Predictions of neural network			Relative error (%)		
1	2485	1065	0.318	2573	982	0.352	3.53	7.79	10.69
2	1852	1173	0.422	1764	1098	0.456	4.75	6.39	8.06
3	1698	1087	0.484	1548	1023	0.439	8.83	5.89	9.29
4	2180	925	0.364	2106	993	0.381	3.39	7.35	4.67
5	2012	856	0.451	2124	928	0.417	5.57	8.41	7.53

* Mean relative error: 6.81(%).

Table 3 The comparison between the real value and prediction of network for (ρ_s, ρ_f, β) in the underlying semi-space

Test groups	Real value (ρ_s, ρ_f, β)			Predictions of neural network			Relative error (%)		
1	2814	1082	0.186	2639	1206	0.163	6.21	11.46	12.37
2	2197	895	0.219	2033	1017	0.236	7.46	13.63	7.76
3	3168	1193	0.204	2983	1109	0.232	5.84	7.04	13.72
4	2513	926	0.172	2785	834	0.190	10.82	9.93	10.46
5	2672	1106	0.197	2410	1183	0.181	9.81	6.96	8.12

* Mean relative error: 9.44(%).

V . CONCLUDING REMARKS

The parameter inverse problem of fluid-saturated porous media is more difficult than that of one phase elastic media. The difficulties come from not only the complicated direct analysis by use of Biot's equations with coupling of solid and fluid phase, but also the extraction of gradient information which is very time-consuming owing to many material constants involved. The mapping method by the neural network reduces an inverse problem to the design of the network structure and the adjustment of weights and value to render the complicated and verbose computation of gradient unnecessary. If only sufficient training sample sets are acquired by direct computation, the inversion mapping from displacement response to material constants can be carried out. A comparison between the first and second examples shows that the mean error in the second example is greater than that in the first. This indicates that larger training sample sets are necessary for more complicated problems in order to obtain satisfactory error limit. In the second example, the prediction of material constants in the overlay is better than that in the underlying semi-space. This implies that the material constants to which the displacement responses are more sensitive can be predicted, in general, to higher precision than that to which the displacement responses are less sensitive.

The generalization of the network is important for the practicability of the neural network trained. The surface of misfit error function in general has more than one extremum. The regularization is to smoothen the surface, which not only improves the prediction of network, but also is helpful for the algorithm to converge to the global extremum. The Tikhonov regularization method and the GCV method for an appropriate selection of regularization parameters are effective. On the other hand, the large number and wide scope of sample sets can improve the prediction of network more directly. But the very large number of sample sets can greatly increase the cost of training. Besides the training algorithm and the training sample sets, the design of structure of network is important to the improvement of the effect of inverse mapping. It's expected to help obtain the best neural network by combines the L-M backpropagation algorithm training network and the evolutionary algorithm for searching for the best structure.

REFERENCES

- [1] Ma, X. R., Tao, L. and Huang, W. H., Inversion method of elastic wave and application, Science Press, Beijing, 1999 (in Chinese).
- [2] Liu, K., Liu, H. W. and Guo, B. Q., Time convolution regularization iteration method for wave equation porosity inversion in 2-D two-phase medium, *Oil Geophysical Prospecting*, Vol.31, 1996, 410 – 414 (in Chinese).
- [3] Liu, K. and Liu, H. W., Method of perturbation for inversion of 2-D elastic wave equation in a two-phase medium, *J. of Harbin University of Architecture and Engineering*, Vol.29, 1996, 80 – 84 (in Chinese).
- [4] Wei, P. J., Zhang, Z. M. and Han, H., Inverse parameter estimation of fluid-saturated porous media by simulated annealing algorithms, *Engineering Mechanics*, Vol.19, 2002, 128 – 131 (in Chinese).
- [5] Wei, P. J., Zhang, Z. M. and Han, H., Parameter inversion of fluid-saturated porous media by genetic algorithms, *Acta Mechanica Solida Sinica*, Vol.23, 2002, 459 – 462 (in Chinese).
- [6] Stavroulakis, G. E. and Antes, H., Crack detection in elastostatics and elastodynamica: a BEM modeling-neural network approach, in: Tanaka, M. and Dulikravich, G. S. (eds), *Inversion Problems in Engineering Mechanics*, Elsevier Science, New York, 1998.
- [7] Feng, X. T., A study on neural network method for displacement inversion analysis, *J. Rock Mech. Engng.*, Vol. 18, 1999, 529 – 533 (in Chinese).
- [8] Sun, D. H., Inversion analysis of mechanics based on neural networks, *Comp. Struc. Mech. Appl.*, Vol. 13, 1996, 308 – 312 (in Chinese).

- [9] Hagan, M. T. and Menhaj, M. B., Training feedforward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks*, Vol.5, 1994, 989 – 993.
- [10] Dominguez, J., Boundary element approach for dynamic poroelastic problems, *Int. J. Numer. Methods in Engng.*, Vol.35, 1992, 307 – 323.
- [11] Gulub, G. C., Heath, M., Wahba, G., Generalization cross-validation as a method for choosing a good ridge parameter, *Technometrics*, Vol.21, 1979, 215 – 223.