

## RESEARCH ON INVERSION OF LIDAR EQUATION BASED ON NEURAL NETWORK

Xingkai Wang<sup>1</sup>, Hu Zhao<sup>2,\*</sup>, Hailun Zhang<sup>2</sup>, Yapeng Liu<sup>1</sup>, Chang Shu<sup>1</sup>

<sup>1</sup>School of computer science and engineering, North MinZu University, Yinchuan, 750021, China - (1395585521, 982843635, 410513982)@qq.com

<sup>2</sup>School of Electrical and Information Engineering, North MinZu University, Yinchuan, 750021, China - zhaoh\_1@yeah.net, 809361497@qq.com

Commission III, WG III/8

**KEY WORDS:** Genetic Algorithm, BP Neural Network, Lidar Equation, Inversion, Extinction Coefficient

### ABSTRACT:

Lidar is an advanced atmospheric and meteorological monitoring instrument. The atmospheric aerosol physical parameters can be acquired through inversion of lidar signals. However, traditional methods of solving lidar equations require many assumptions and cannot get accurate analytical solutions. In order to solve this problem, a method of inverting lidar equation using artificial neural network is proposed. This method is based on BP (Back Propagation) artificial neural network, the weights and thresholds of BP artificial neural network is optimized by Genetic Algorithm. The lidar equation inversion prediction model is established. The actual lidar detection signals are inverted using this method, and the results are compared with the traditional method. The result shows that the extinction coefficient and backscattering coefficient inverted by the GA-based BP neural network model are accurate than that inverted by traditional method, the relative error is below 4%. This method can solve the problem of complicated calculation process, as while as providing a new method for the inversion of lidar equations.

### 1. INTRODUCTION

Air pollution has always been a concern problem for human beings. With the development of industrialization, atmospheric aerosols are increased continually, which has a great impact on climate, environment and agriculture. Atmospheric aerosols have numerous natural and artificial sources, such as volcanic eruptions, sea spray, ground dust, combustion of organisms, human activities, and use of fuels, all of them produce various particles (Chen et al. 2011, Bryukhanova and Abramotchkin, 2000). In recent years, researchers pay more attention to the detection of atmospheric aerosols. Lidar is an advanced instrument for the detection and monitoring of atmospheric aerosols. Through the inversion of the lidar signals, we can obtain the atmospheric aerosol microphysical parameters (Shen, 1999). The uncertainty of the inversion of lidar equation has always been a problem, because the lidar equation is a ill-posed equation, it has two unknowns (extinction coefficient and backscattering coefficient). In order to solve the two unknowns, the lidar ratio (extinction coefficient/backscattering coefficient) must be set. The lidar ratio is usually determined by experience, so it always makes the solution of lidar equation has high degree of uncertainty (Zhao et al. 2018).

At present, there are three traditional methods to solve the lidar equation: Collis method, Klett method and Fernald method. Collis slope method assumes that the atmosphere is uniformly distributed. However, the reality of non-uniform distribution of atmospheric aerosols cannot be ignored, so this method will lead to large errors. Klett (Klett, 1985) proposed a single-component fitting method, although it could overcome the limitation of uniform atmosphere and make the results more practical. However, when the extinction effect of aerosols and atmosphere has little difference, the extinction effect of atmosphere cannot

be ignored. This method is no longer applicable. At present, Fernald method is widely used (Fernald, 1984). Fernald method regards the atmosphere as two parts: air molecules and aerosols. However, the solution solved by Fernald method is greatly related to the actual atmospheric state, and the solution is highly uncertain.

BP neural network is a multilayer feed forward neural network (Wen et al. 2000), which has a flexible structure and can learn the nonlinear relationship between input and output excellently. By mastering this rule, the result close to the expected output can be obtained according to the input.

Currently, Li et al. (2018) proposed An RBF neural network approach for retrieving atmospheric extinction coefficients based on lidar measurements. However, when the training sample of RBF neural network is large, the number of hidden layer neurons is much larger than that of BP neural network, and the complexity of RBF neural network also increases greatly. BP neural network are optimized by the genetic algorithm can solve the problem of BP neural network easily trapped in local extremes (Ding et al. 2011). It has more advantages on the accuracy of convergence, than RBF. Therefore, in this paper, the genetic algorithm is used to optimize the BP neural network model to invert extinction coefficient, the echo power as the training input samples, and the existing extinction coefficient as the training output samples. Through GA-BP self-learning, the mapping relationship between the echo power and the extinction coefficient is learned, and inversion prediction effect is obtained.

\* Corresponding author: Hu Zhao - zhaoh\_1@yeah.net

## 2. BP NEURAL NETWORK BASED ON GENETIC ALGORITHM

### 2.1 BP Neural Network

In this experiment, the relationship between lidar echo power and atmospheric extinction coefficient is nonlinear. Artificial neural network is a mathematical model or computational model which is designed to simulate human brain neural network. It simulates human brain neural network in terms of structure, implementation mechanism and function. The trained network model can calculate complex mathematical relations (Seyed, 2013), which is exactly suitable for the inversion and prediction of lidar equations.

The BP neural network generally refers to the multilayer feed forward neural network trained by BP (Error back propagation) algorithm, which is the most successful neural network learning algorithm so far (Yann et al. 1989).

BP neural network consists of three parts: input layer, hidden layer and output layer. In this experiment, the input layer data is the laser radar echo power, and the output layer data is the extinction coefficient value. Each layer has a number of neurons, which are fully connected in the adjacent layer, but not connected in the same layer. Neurons are topological network established according to the activity mechanism of human brain, as shown in Figure 1.

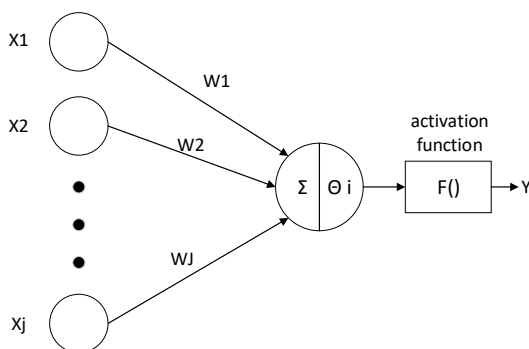


Figure 1. Neurons

In the figure,  $X_i$  as the input value,  $W_i$  as the weight,  $\Theta_i$  as the threshold value,  $Y$  as the output value.

The characteristic of BP neural network is: its signal is forward propagation, and the error is back propagation. This characteristic enables BP neural network to continuously adjust and optimize itself according to the error value. When the error between the predicted output extinction coefficient value and the expected output value cannot meet the specified requirements, the output layer will transmit the error to all neurons in each layer in a certain form. The error is the basis for the neuron to modify the weight value, so that the predicted output extinction coefficient value is close to the expected output value until we get the expected extinction coefficient value or the number of iterations reached has to stop.

### 2.2 Genetic Algorithm

Genetic algorithm is a parallel search algorithm used to solve global optimization problems. It comes from Darwin's biological theory of evolution. The selection, crossover and mutation on

chromosome are operated through the simulation of the biological evolution process. The individuals with better fitness will be left, and the bad ones will be discarded according to the fitness function that has been written. This is just like the rule of survival of the fittest in nature, which helps us better improve the training accuracy and get more accurate extinction coefficient through inversion.

### 2.3 GABP

Although the learning ability and generalization ability of BP neural network are excellent, it is easy to fall into local extremum. It will affect the prediction results, and the slow convergence speed will make the network become inefficient. As a global optimization algorithm, genetic algorithm can reduce the risk of falling into local extremum, and parallel search can make the network more efficient. Therefore, the more accurate extinction coefficient can be obtained by optimizing the weights and thresholds of BP neural network by genetic algorithm.

The process of genetic algorithm to optimize BP neural network mainly includes three parts: determination of BP neural network connection structure, optimization of BP neural network weights and thresholds by genetic algorithm, and prediction of BP neural network (Wang and Cai, 2003, Huang et al. 2009). The algorithm process is shown in Figure 2.

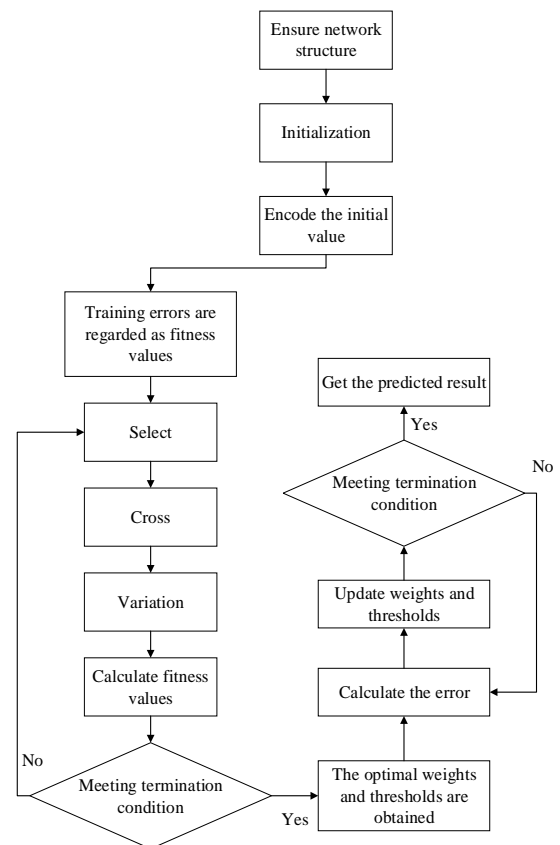


Figure 2. Algorithm process

## 3. EXPERIMENTAL DESIGN

### 3.1 Data Processing

The lidar equation (Klett, 1981) is known as equation 1.

$$p(r) = p_0 Y(r) \frac{c t_p}{2} g \frac{A_R}{r^2} \beta(r) T^2(r) \quad (1)$$

Where,  $p(r)$  is the echo power at the range  $r$ , it can be got from an oscilloscope.  $Y(r)$  is geometric overlap factor,  $\beta(r)$  is backscattering coefficient,  $T(r)$  is the atmospheric transmittance,  $T(r) = e^{-\int_0^r \sigma(r') dr'}$ , where  $\sigma(r')$  is the atmospheric extinction coefficient.

$\beta(r)$  and  $\sigma(r')$  is unknown in the lidar equation. You just have to solve for one, and you can solve for the other very easily. Therefore, we take echo power  $p(r)$  as input data and extinction coefficient  $\sigma(r')$  as output data.

First, the original sets of data in the case of cloud weather are obtained through the Klett method. Since the collected data are invalid at the first 1036 points, the data range is from 54 meters to 12,000 meters, 8,000 points are selected. The input data and output data are 8000 groups respectively. Next, in order to make the training results more excellent, we randomly selected 2000 groups of input data from 8000 groups as training samples, and took all 8000 groups of input data as test samples. The output data were processed in the same way as the input data.

After data grouping, in order to ensure the effectiveness of data in the training process, the data must be normalized. It can improve the convergence speed of the network. In the experiment, `mapminmax` function in Matlab software is used to normalize the echo power and extinction coefficient values. The normalized formula is:

$$X_k = (X_k - X_{\min}) / (X_{\max} - X_{\min}) \quad (2)$$

Where,  $X_{\min}$  is the minimum number in the data series, and  $X_{\max}$  is the maximum number in the data series.

### 3.2 Hidden Layer Neuron

The selection of the number of hidden layer neurons has a great influence on the experimental results. Over fitting may occur if the selection of the final value is too large.

In the experiment, there was a phenomenon of over fitting. When the trained network model is use to predict the extinction coefficient, the relative error obtained from the training is very low, but when the randomly selected from 1000 groups of echo power to predict the extinction coefficient, We compared the predicted extinction coefficient with that obtained by Klett method, it is found that there was a big difference.. This is the phenomenon of over fitting. Then we solved the over-fitting phenomenon by reducing the number of hidden layer neurons and training precision.

For the selection of the number of neurons in the hidden layer, generally the approximate range is calculated using the empirical formula. Then the optimal value is selected through repeated experiments. If the selection of the final value is too large, overfitting phenomenon may occur. If it is selected too small, the prediction error may be too large. The empirical formula is as follows:

$$L = \sqrt{n+m+a} \quad (3)$$

Where,  $m$  is the number of neurons in the output layer,  $L$  is the number of neurons in the hidden layer,  $n$  is the number of neurons in the input layer, and  $a$  is a constant and  $0 < a < 10$ .

Because the input data and the output data are 1, according to the empirical formula, we can get the range of the number of neurons in the hidden layer is [1,11]. Through repeated experiments, the number of neurons in the hidden layer is set as 5.

After knowing the number of neurons in each layer, The BP neural network structure of the experiment can be obtained, which is shown in Figure 3.

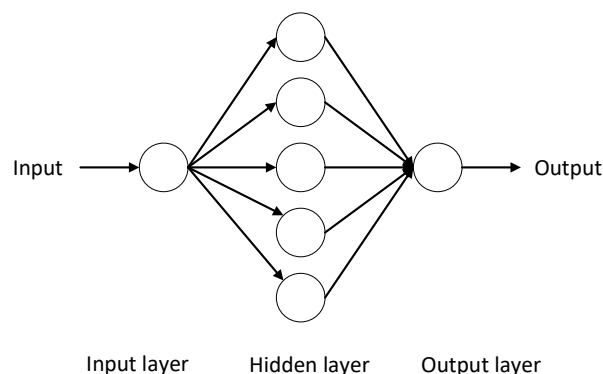


Figure 3. Structure of neural network

### 3.3 Parameter Setting

The parameters setting of BP neural network optimized by genetic algorithm in this experiment is shown in Table 1.

Name of parameter	Parameter setting
Evolution algebra	10
The training accuracy	0.0001
Number of learning iterations	500
Hidden layer activation function	Tansig
Output layer activation function	Purelin
The training function	Trainlm
Crossover probability	0.5
Mutation probability	0.1
Population size	30

Table 1. Parameter settings

## 4. THE EXPERIMENTAL RESULTS

The neural network model is used to train and predict the processed data. The error judgment of this experiment adopts the method of relative error, and the calculation formula is shown in equation 4.

$$\text{errors} = (\text{test\_simu} - \text{output\_test}) / \text{output\_test} \quad (4)$$

In the end, the relative error was 2.04%. Several experiments were carried out, and the relative error was controlled below 4 %.

Figure 4 shows the extinction coefficient predicted by GA-BP neural network and inverted by Klett method.

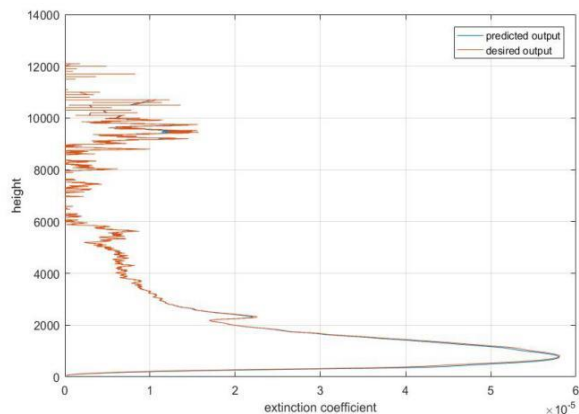


Figure 4. Comparison of predictions

In Figure 4, the blue line represents the extinction coefficient predicted by GA-BP neural network, and the red line represents the extinction coefficient retrieved by Klett method. It can be seen from the figure that the results obtained by the two methods are very close, which proves that the training effect of this experiment is very good. In order to better observe and predict the results, the partial enlarged figure is shown in Figure 5.

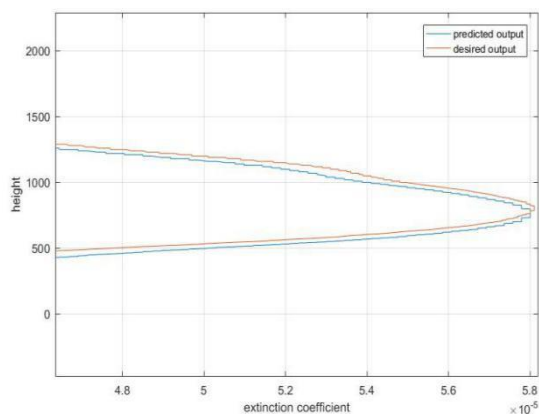


Figure 5. Partial enlarged view

Figure 6 is a regression diagram, it can be seen that the results of training, verification and testing of neural network. If the data fit perfectly, all the data in the four figures should be on the diagonal of the 45° Angle. If the deviation is large or the data points are too few, it means that the training results may be over fitted or the number of training data is too small. It can be seen from the figure that the data fitting effect is good, the predicted output and expected output are basically the same, and the R value of the fitting output is greater than 0.99.

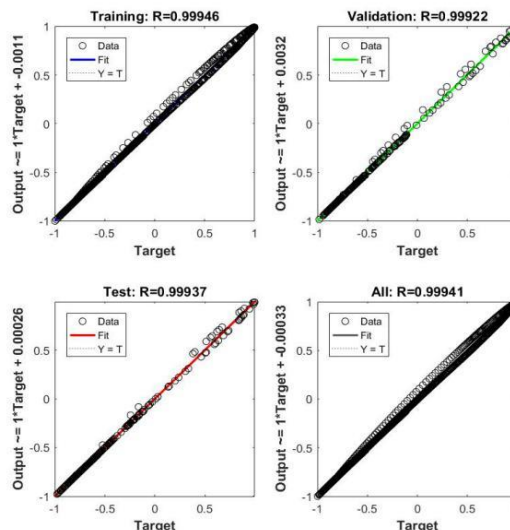


Figure 6. Regression diagram

In order to verify the reliability of training samples of 2000 sets of data, 2000 groups of data, 3000 groups of data and 4000 groups of data are selected, in this paper, as training samples respectively. The final relative error is used to judge it. The variation trend of the error is shown in Figure 7.

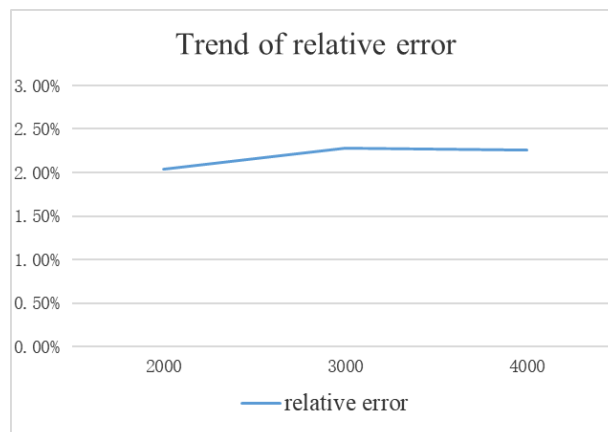


Figure 7. Error changes

It can be seen from the Figure 7 that the training error does not change much when the training sample selection is 2000 group, 3000 group and 4000 group. The training error is 2.04%, 2.28% and 2.26% respectively. The training error is minimum when training sample selection 2000 group. So it can be concluded that the training effect has little relation with the number of training samples. This does not include the underfitting caused by too few training samples.

The above is the case that the extinction coefficient is predicted in the cloud weather. In order to ensure the feasibility, further experiments are carried out in no cloud, and the experimental results are shown in Figure 8 and 9.

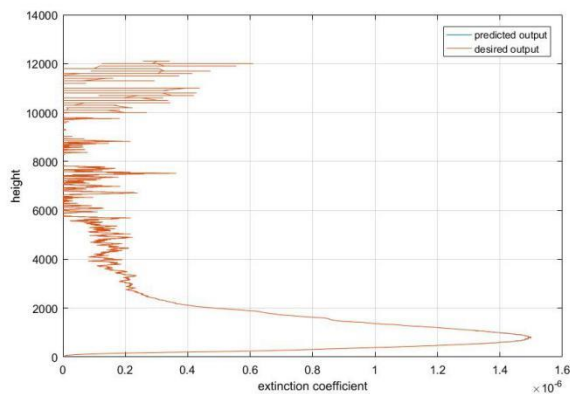


Figure 8. Prediction of extinction coefficient under cloudless conditions

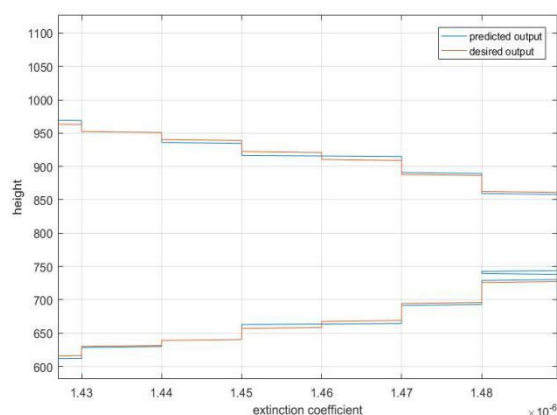


Figure 9. Partial enlargement under cloudless conditions

The training effect is still excellent with a relative error of 0.2% under cloudless conditions. The extinction coefficient obtained through GA-BP neural network model is very consistent with that obtained by Klett method.

## 5. CONCLUSION

The BP neural network optimized by the genetic algorithm was used to invert the lidar equation. The relative error was 2.04 % in the case of cloud and 0.2% in the case of cloudless, it is identified with the extinction coefficient obtained by Klett method. Through repeated experiments and verification, it can be concluded that it is effective using the GA - BP inversion method to predict the lidar equation. it can get more accurate extinction coefficient though optimizing the parameters by GA-BP repeatedly to. In addition, there is no need to carry out a lot of calculations and assumptions like traditional mathematical methods. GA-BP model can save a lot of time for research and make the research results more accurate.

In the future work, we will continue to improve the GA-BP inversion method, looking for ways to improve the algorithm, make the results more accurate. On this basis, more different types of neural network models will be tried and applied to the inversion of atmospheric parameters.

## ACKNOWLEDGEMENTS

This research was funded by the National Natural Science Foundation of China (Grant No. 61865001) and the Ningxia

Natural Science Foundation (Grant No. 2018AAC03103).

## REFERENCES

- Bryukhanova, V. V., Abramotchkin, S. A., 2000. A lidar for large-scale droplet aerosol particle sensing in the atmosphere. *Modern Techniques & Technology, Mtt VI International Scientific & Practical Conference of Students, Post-graduates & Young Scientists. IEEE*, 94-97. 10.1109/SPCMTT.2000.896063.
- Cun, Y. L., Boser, B., Denker, J. S., Howard, R. E., Habbard, W., Jackel, L. D., et al. 1990. Handwritten digit recognition with a back-propagation network. *Advances in Neural Information Processing Systems*, 2(2), 396-404.
- Dongya, Shen., 1999. Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: simulation. *Appl Opt*, 38(12), 2346-2357. 10.1364/AO.39.001879.
- Ding, S., Su, C., Yu, J., 2011. An optimizing BP neural network algorithm based on genetic algorithm. *Artificial Intelligence Review*, 36(2), 153-162. 10.1007/s10462-011-9208-z.
- Fernald, F. G., 1984. Analysis of atmospheric lidar observations : some comments. *Applied Optics*, 23(5), 652-653. 10.1364/AO.23.000652.
- Hu, Z., Jiandong, M., Chunyan, Z., Xin, G., 2018. A method of determining multi-wavelength lidar ratios combining aerodynamic particle sizer spectrometer and sun-photometer. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 217, 224-228. 10.1016/j.jqsrt.2018.05.030.
- Hongxu, L., Jianhua, C., Fan, X., Binggang, L., Zhenxing, L., Lingyan, Z., et al., 2018. An RBF neural network approach for retrieving atmospheric extinction coefficients based on lidar measurements. *Applied Physics B*, 124(9), 184-. 10.1007/s00340-018-7055-1.
- Jian-Guo, H., Hang, L., Hou-Jun, W., Bing, A. L., 2009. Prediction of time sequence based on ga-bp neural net. *Journal of Electronic Science and Technology of China*, 38(5), 687-692. 10.3969/j.issn.1001-0548.2009.05.028.
- Jin, W., Li, Z. J., Wei, L. S., Zhen, H., 2000. The improvements of BP neural network learning algorithm. *Signal Processing Proceedings, 2000. WCCC-ICSP 2000. 5th International Conference on. IEEE*, 1647-1649. 10.1109/ICOSP.2000.893417.
- Klett, J. D., 1985. Lidar inversion with variable backscatter/extinction ratios. *Applied Optics*, 24(11), 1638-1643. 10.1364/AO.24.001638.
- Klett JD., 1981. Stable analytical inversion solution for processing lidar returns. *Applied Optics*, 20(2), 211-220. 10.1364/AO.20.000211.
- Shengzhe, Chen., Yinchao, Zhang., Siying, Chen., He, Chen., 2011. Comparing methods for retrieving aerosol extinction coefficient with U.S. Standard Atmospheric Model and Temperature Gradient. *International Conference on Remote Sensing. IEEE*, 2043-2046. 10.1109/RSETE.2011.5964706.
- Saeed Madani, S., 2013. Electric load forecasting using an artificial neural network. *IEEE Transactions on Power Systems*, 6(2), 442-449. 10.1109/59.76685.

Wang, S., Cai, J., 2003. Application of hybrid algorithm based on ga-bp in transformer diagnosis using gas chromatographic method. *High Voltage Engineering*, 29(7), 3-6.