Vol 18 No 6, June 2009
1674-1056/2009/18(06)/2441-04

Chinese Physics B

© 2009 Chin. Phys. Soc.
and IOP Publishing Ltd

Prediction of the plasma distribution using an artificial neural network*

Li Wei(李 炜)[a], Chen Jun-Fang(陈俊芳)[a]†, and Wang Teng(王 腾)[b]

[a] *School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou* 510006, *China*

[b] *School of Computer, South China Normal University, Guangzhou* 510631, *China*

(Received 2 December 2008; revised manuscript received 19 December 2008)

In this work, an artificial neural network (ANN) model is established using a back-propagation training algorithm in order to predict the plasma spatial distribution in an electron cyclotron resonance (ECR) — plasma-enhanced chemical vapor deposition (PECVD) plasma system. In our model, there are three layers: the input layer, the hidden layer and the output layer. The input layer is composed of five neurons: the radial position, the axial position, the gas pressure, the microwave power and the magnet coil current. The output layer is our target output neuron: the plasma density. The accuracy of our prediction is tested with the experimental data obtained by a Langmuir probe, and ANN results show a good agreement with the experimental data. It is concluded that ANN is a useful tool in dealing with some nonlinear problems of the plasma spatial distribution.

Keywords: artificial neural network, ECR-PECVD plasma, distribution
PACC: 5270D, 5225L, 0540

1. Introduction

ECR plasma technology has been widely used in plasma processing such as film deposition, plasma etching and material surface treatment, because it has many good characteristics such as low gas pressure and high plasma density. During recent decades, there has been a lot of research on the application of the ECR plasma system.[1−3] Although the spatial distribution of ECR plasma is very important in plasma processing technology, it was less investigated, and the scope of investigated data was limited by the cost and time.[4−7] Therefore, it is necessary to obtain more desirable parameters by other means.

ANN is a massive parallel array of computational nonlinear units that models some of the functionality of the human brain and attempts to capture some of its computational strength. It is highly effective to learn the technique that is suitable for nonlinear complex tasks with a high degree of accuracy. Recently, some researchers have attempted to use ANN models for various applications in different fields.[8−20]

In this paper, an ANN model is studied to predict the nonlinear spatial distribution of ECR-PECVD plasma.

2. Experimental method

A schematic diagram of the ECR-PECVD plasma system is shown in Fig.1. It consists of a waveguide 1, a resonance room with dimensions ϕ14×20 cm 2, a magnetic coil 3, a reaction chamber with dimensions ϕ35×50 cm 4, a substrate table 5, a vacuum system 6, and a mass flow controller 7. Variation of magnetic field strength can be obtained by adjusting the current in the magnetic coil (I_m). The working gas N_2 is introduced into the resonance chamber by the mass flow controller (MFC); and the energy

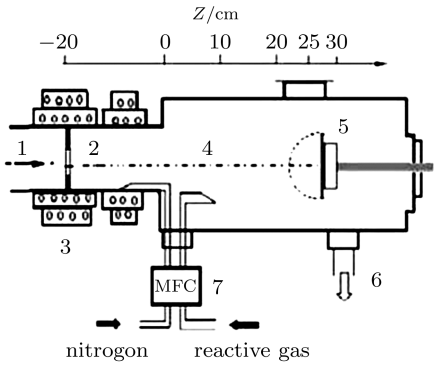


Fig.1. The schematic diagram of the ECR-PECVD plasma system.

*Project supported by the National Natural Science Foundation of China (Grant No 10575039).
†Corresponding author. E-mail: tolwwt@163.com

of the microwave is strongly absorbed by the plasma in the ECR zone. As a result, high-density plasma is generated with a high-degree of ionization.

The spatial distribution of plasma density in the ECR-PECVD plasma system is measured by a Langmuir probe. The plasma spatial distribution is mainly determined by five factors: the radial position, the axial position, the gas pressure, the microwave power and the magnetic fields. So, we make our measurements at various radial positions (R = 0, 2, 4, 6, 8, 10, 12, 14 cm), axial positions (Z = 0, 5, 10, 15, 20, 25 cm), gas pressure (P_1 = 0.03, 0.05, 0.07, 0.09, 0.10, 0.12, 0.14, 0.16 Pa), microwave power (P_2 = 300, 400, 500, 600, 650, 800 W) and magnet coil currents (I_m = 140, 145, 150, 155, 160, 165 A). Three hundred of the data points from above are randomly chosen as the training set.

3. The ANN model

ANN is a computational model that is mainly composed of three layers of neurons: the input layer, the hidden layer and the output layer. The units of the system are connected by the weights. The input layer consists of all the input factors, the information from which is then processed through the hidden layer, and the resulting information is computed in the output layer. The structure of an ANN system can be shown schematically in Fig.2.

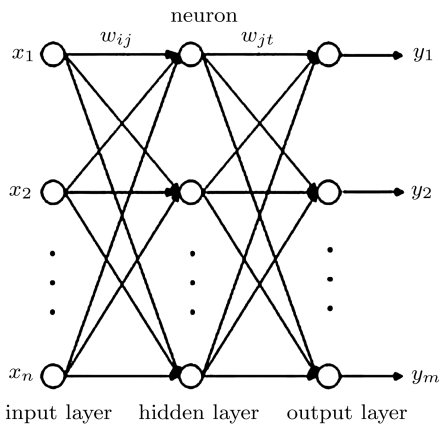


Fig.2. The schematic structure of an ANN system.

There are several kinds of training algorithm available in the ANN system. In this study, a back-propagation (BP) algorithm is used to train a three-layer feed forward neural network. Key steps for establishment of the model can be summarized as follows:

a) Initialize the values of the connection weights w_{ij} (between the neurons in the input layer and the hidden layer) and v_{jt} (between the neurons in the hidden layer and the output layer), the output threshold values θ_j (of the neurons in the hidden layer) and γ_k (of the neurons in the output layer), which are random values in the interval $[-1, 1]$.

b) Offer the input data $P_k = (a_1^k,\, a_2^k, \ldots, a_n^k)$ to the neurons of the input layer.

c) Calculate the input s_j and the output b_j of neurons in the hidden layer by using Eqs.(1) and (2), where $f(s_j)$ is a transfer function,

$$s_j = \sum_{i=1}^{n} w_{ij} a_i - \theta_j, \tag{1}$$

$$b_j = f\left(s_j\right). \tag{2}$$

d) Calculate the input L_k and the output C_k of neurons in the output layer by using Eqs.(3) and (4), where $f(L_k)$ is a transfer function,

$$L_k = \sum_{j=1}^{p} w_{jk} b_j - \gamma_k, \tag{3}$$

$$C_k = f\left(L_k\right). \tag{4}$$

e) Calculate the correction factor d_t for all units in the output layer using the output C_k and the target output $T_k(y_1^k,\, y_2^k, \ldots, y_n^k)$.

$$d_t = \left(y_t - C_t\right) \cdot C_t \cdot \left(1 - C_t\right). \tag{5}$$

f) Correct the connection weights v_{jt} and the output threshold γ_k,

$$v_{jt}^{\text{new}} = v_{jt}^{\text{old}} + \alpha \cdot d_t \cdot b_j, \tag{6}$$

$$\gamma_t^{\text{new}} = \gamma_t^{\text{old}} + \alpha \cdot d_t, \tag{7}$$

where α is the learning rate.

g) Correct the connection weights w_{ij} and the output threshold θ_j by using equations similar to Eqs.(6) and (7).

h) Return to the second step and repeat with a new set of input values.

The iteration continues until the overall error between calculated and target output approaches the pre-set error criteria.

In our ANN model, there are three layers: the input layer, the hidden layer and the output layer. The input layer is composed of five neurons: the radial position, the axial position, the gas pressure, the microwave power and the magnet coil current. The numbers of the neurons in the hidden layer are 3, 10

and 12. The output layer is our target output neuron: the plasma density. So, the structures of our ANN model are 5-3-1, 5-10-1 and 5-12-1.

The neural network is realized in MATLAB7.0 on the platform of Microsoft Windows XP. The neural network requires that the range of both the input values and the output values should be between 0 and 1, so, all data in our model are unified.

4. Results and discussion

The effects of training cycles on the network performance with different structures are shown in Fig.3, in which the vertical axis indicates the difference between the training performance and the training goal. We set the training goal to 1×10^{-4} to simplify the model.

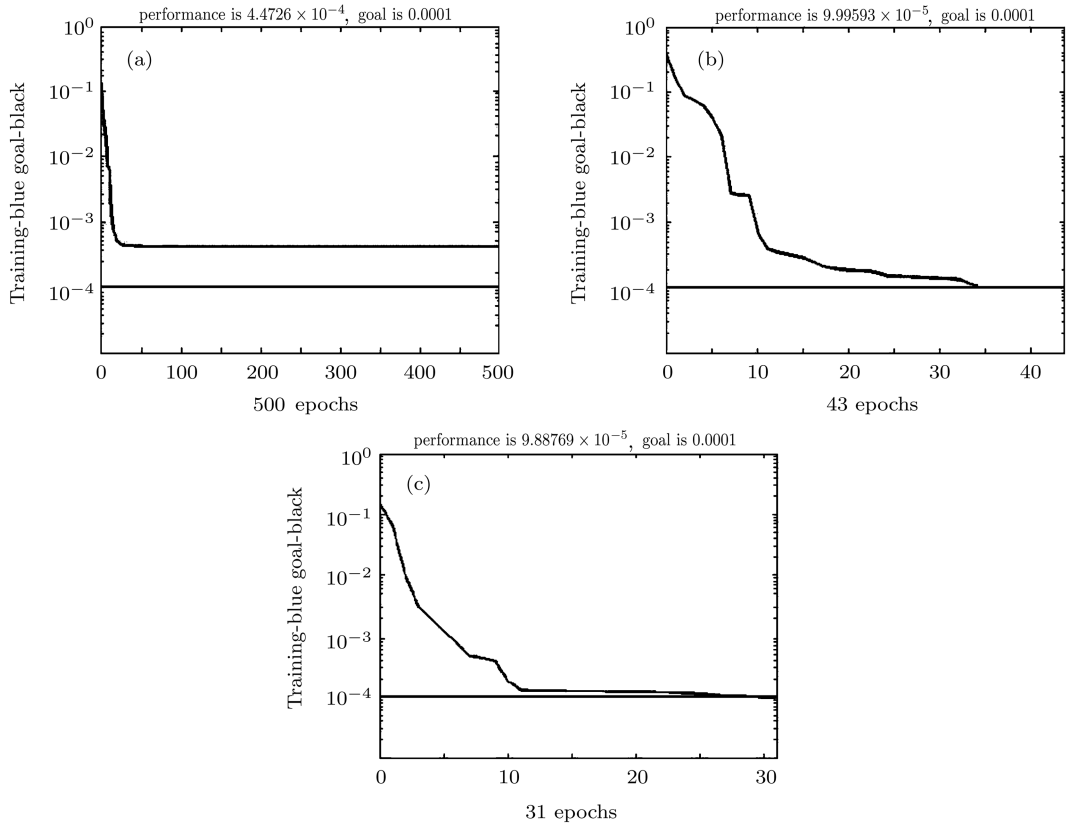


Fig.3. The effect of the training cycles on the different network performance in different structures, (a) 5-3-1; (b) 5-10-1; (c) 5-12-1.

From Fig.3 it can be seen that the ANN with the structure of 5-3-1 cannot attain our training goal, and only the ANN with the structure of 5-12-1 is the most efficient design. In this 5-12-1 ANN model, when the neural network is trained up to 30 times, the goal of training accuracy can be attained. It is clear from Fig.3 that further training cycles have no considerable effect on the network performance. The errors of the training set decrease dramatically at the beginning of the training cycles, then gradually tend to approach the minimum value.

The accuracy of prediction in the ANN model with different structures is shown in Fig.4. The vertical axis indicates the error between six random experimental data from the training set and corresponding tested data from two different ANN models. The error bar was chosen as 0.005. The horizontal axis indicates the serial numbers of those experimental data. From this figure, we find most of data were inside the error bar, so they are believable. But some data of 5-10-1 ANN model were outside of the error bar. From this figure, we can conclude that the 5-12-1 ANN model has a better ability to predict than the 5-10-1 ANN model. It is believed that the number of neurons in the hidden layer plays an important role.

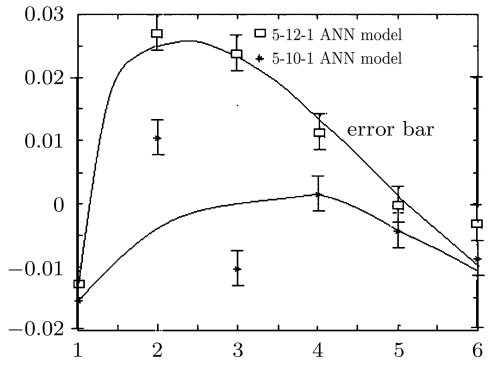


Fig.4. Comparisons of errors in predicting data using the ANN model with different structures.

In order to make the model more effective, we again choose some predicted data that are not same as the data in the training set and measured the predicted data again. Corresponding experimental data of the test set in the 5-12-1 ANN model are given in Table 1. It is shown that the predicted and experimental values are very close to each other. The relative error is less than 6.5%. It can be concluded that the values of the plasma density predicted by the 5-12-1 ANN model are in good agreement with the experimental data. The ability of the ANN model to accurately predict the spatial distribution in an ECR-PECVD system can be concluded.

Table 1. Data used in the prediction and the predicted values.

gas pressure	microwave power	magnet coil current	axial position	radial position	plasma density/cm^3		
/Pa	/W	/A	/cm	/cm	experimental data	predicted data	relative error
0.07	650	145	0	8	2.260	2.126	5.9%
0.10	650	145	10	4	5.406	5.529	2.3%
0.12	400	150	10	14	1.985	1.863	6.1%
0.10	500	145	15	8	3.824	3.854	0.7%
0.09	600	155	20	10	3.405	3.365	1.1%
0.14	600	160	30	2	2.284	2.189	4.2%

5. Conclusions

In the present work, a relative new technique, ANN associated with nonlinear systems, is used in modeling the prediction of the spatial distribution in an ECR-PECVD plasma system. The effectiveness of the ANN model is demonstrated, and the results obtained using this model show a good agreement with existing experimental data. It can be concluded that ANN is a useful tool in dealing with some complex nonlinear problems of plasma spatial distribution and that using ANN can considerably save cost and time.

References

[1] Mayuko Koga and Hiroshi Muta 2006 *Vacuum* **80** 771
[2] Chen J F and Ren Z X 1999 *Vacuum* **52** 411
[3] Toader E I 2004 *Plasma Sources Sci. Technol.* **13** 646
[4] Musil J 1996 *Vacuum* **47** 145
[5] Hiroshi M, Doan H T and Yoshinobu K 2005 *Surf. Coat. Technol.* **200** 850
[6] Wang L, Cao J X, Wang Y, Niu T Y, Wang K and Zhu Y 2007 *Acta Phys. Sin.* **56** 1429 (in Chinese)
[7] Jin X Y, Qiu X J and Zhu Z Y 2006 *Acta Phys. Sin.* **55** 5338 (in Chinese)
[8] Wang Y S, Sun J, Wang C J and Fan H D 2008 *Acta Phys. Sin.* **57** 6120 (in Chinese)
[9] Wang H, Wang A K, Yang Q W, Ding X T, Dong J Q, Sanuki H and Itoh K 2007 *Chin. Phys.* **16** 3738
[10] Wu H M, Graves D B and Porteos R K 1995 *Plasma Sources Sci. Technol.* **4** 22
[11] Dorteoust, Wu H M and Graves D B 1994 *Plasma Sources Sci. Technol.* **3** 25
[12] Lungu C P and Iwasak K 2002 *Vacuum* **66** 197
[13] Du X L and Chen G C 1999 *Acta Phys. Sin.* **48** 257 (in Chinese)
[14] Z Sterjovski and D Nolan 2006 *J. Mater. Proc. Technol.* **170** 536
[15] Xu L J and Xing J D 2007 *Mater. Des.* **28** 1425
[16] Bezerra E M and Ancelotti A C 2007 *Mater. Sci. Eng.* A **464** 177
[17] Scott D J and Coveney P V 2007 *J. Eur. Cerma. Soc.* **27** 4425
[18] Ming D, Jean, Chyuan D L and Wang J T 2005 *Appl. Surf. Sci.* **245** 290
[19] Zhang G and Guessasma S 2006 *Surf. Coat. Technol.* **200** 2610
[20] Hakan C and Hasan O 2006 *Wear* **261** 064