Characterisation of the plasma density with two artificial neural network models*

Wang Teng(王 腾)[a)b)], Gao Xiang-Dong(高向东)[a)†], and Li Wei(李 炜)[c)]

[a)] *Faculty of Electromechanical Engineering, Guangdong University of Technology, Guangzhou* 510090, *China*
[b)] *School of Computer, South China Normal University, Guangzhou* 510631, *China*
[c)] *School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou* 510006, *China*

(Received 25 April 2009; revised manuscript received 4 August 2009)

This paper establishes two artificial neural network models by using a multi layer perceptron algorithm and radial based function algorithm in order to predict the plasma density in a plasma system. In this model, 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 the target output neuron: the plasma density. The accuracy of prediction is tested with the experimental data obtained by the Langmuir probe. The effectiveness of two artificial neural network models are demonstrated, the results show good agreements with corresponding experimental data. The ability of the artificial neural network model to predict the plasma density accurately in an electron cyclotron resonance-plasma enhanced chemical vapour deposition system can be concluded, and the radial based function is more suitable than the multi layer perceptron in this work.

Keywords: plasma density, prediction, multi layer perceptron, radial based function
PACC: 0540, 5270D, 5225L

1. Introduction

The technology of low-temperature plasma has been widely used in the fields of microelectronics, materials, chemical industry, film deposition and plasma etching. It is very important to study inner physical parameters in order to optimise the technology process. During the recent decades, there has been a lot of research on the application of the electron cyclotron resonance (ECR) plasma system,[1−3] because the ECR plasma system has many good characteristics such as low gas pressure and high plasma density. Although the plasma density of ECR plasma is very important to the plasma processing technology, it was less investigated, and the scope of investigating data was limited by the cost and time.[4−7] So, it is necessary to obtain more desired parameters from some other means.

Artificial neural network (ANN) is now already a very well developed, natural and routine technique, which utilises interconnected nodes or neurons to form a network that can model a complex functional relationship. The technique is particularly suited to problems which involve the manipulation of multiple parameters and non-linear interpolation. Recently, some researchers attempted to use ANN models for various applications in various fields.[8−20] But there is little information about the application of ANN in the simulation of plasma density distribution.

In our work, two ANN models are studied to characterise the nonlinear plasma density of ECR-plasma enhanced chemical vapour deposition (PECVD) plasma system.

2. The experimental methods

The schematic diagram of the ECR-PECVD plasma system is showed in Fig. 1. It consists of a wave-guide (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 (MFC) (7). The various magnetic field strengths can be obtained by adjusting the magnetic coil currents ($I_{\rm m}$). The working gas N_2 is introduced into the resonance chamber by the MFC. The energy of the microwave is strongly absorbed by the plasma in the ECR zone. As a result, high-density plasma is generated with high-degree of ionisation.

*Project supported by the National Natural Science Foundation of China (Grant No. 60375012).
†Corresponding author. E-mail: gaoxd666@126.com
© 2010 Chinese Physical Society and IOP Publishing Ltd

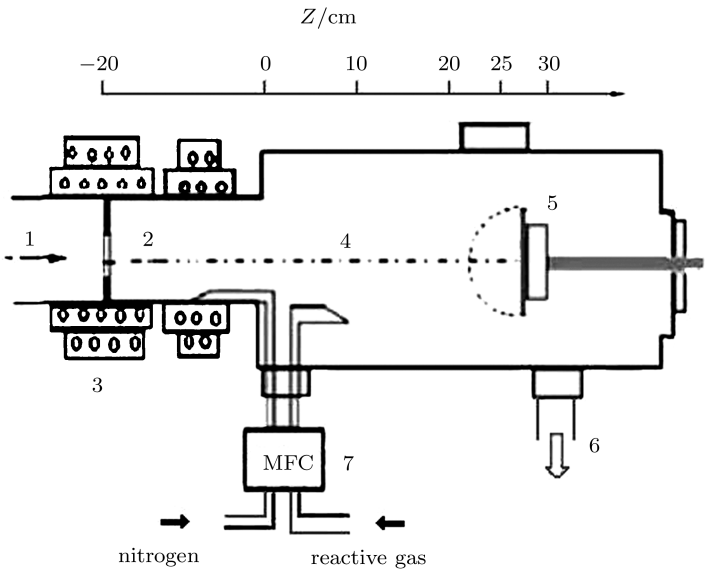


Fig. 1. The schematic diagram of the ECR-PECVD plasma system.

The plasma density in the ECR-PECVD plasma system is measured by a Langmuir probe. The plasma density 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 ($Z = 0$, 10, 15, 20, 25, 30 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 ($I_{\rm m} = 140$, 145, 150, 155, 160, 165 A). Three hundreds data from above are randomly chosen as the training set.

3. The ANN model

There are several kinds of training algorithm available in the ANN system. In this study, a multi layer perceptron (MLP) algorithm and a radial based function (RBF) algorithm are used to train the artificial neural network.

The MLP networks are 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. Information from the input layer is then processed through the hidden layer, and the following information is computed in the output layer.

The MLP training algorithm is an iterative gradient algorithm, designed to minimise the mean square error between the predicted output and the actual output. The key steps for establishment of the MLP model can be summarised briefly.

Initialisation: prior to the first iteration of loop, initialise the values of the connection weights and the output threshold values of the neurons, which have random values in [–1, 1].

Maintenance: offer the input data to the neurons of the input layer. Then calculate the input and the output of neurons in the hidden layer and in the output layer. After each loop, the error between the predicted output and the actual output are propagated backward to adjust the weight in a manner mathematically guaranteed to converge.

Termination: the iteration continues until the overall error between calculated and target output is approaching to the pre-set error criteria.

The RBF networks have a very strong mathematical foundation rooted in regularisation theory for solving ill-conditioned problems. As its name implies, radially symmetric basis function is used as activation functions of nodes in hidden layer. The transformation from the input layer to the output $f(x)$ is computed by

$$f(x) = \sum_{i=1}^{n} W_i R_i(x) + W_0,$$

where W_i $(i = 1, 2, \ldots, n)$ is the connection weight between the hidden layer and output layer, W_0 is the

bias, $R_i(x)$ $(i = 1, 2, \ldots, n)$ are RBFs,

$$R_i(x) = \varphi \left\| X - C_i \right\|,$$

$\varphi \left\| X - C_i \right\|$ is the activation function in the jargon of the ANN, here it is a Gaussian function,

$$R_i = -\exp\left(-\sum_{j=1}^{n} \frac{\left\| X_i - C_i \right\|^2}{2\sigma_{ij}^2} \right),$$

where $C_i^{\mathrm{T}} = [C_{i1}, C_{i2}, C_{i3}, \ldots, C_{in}]$ $(i = 1, 2, \ldots, n)$ is the centre of the receptive field, and σ_{ij} $(i, j = 1, 2, \ldots, n)$ is the width of the Gaussian function, which indicates the selectivity of the neuron. The centre C can be randomly chosen by the simplest way from the train set. Another approach is to use the k-means technique of clustering input training set into groups and choose the centre of each group as the centre. Also, C can be treated as a network parameter along with W_i and adjusted through error-correction training. After the centre C is determined, the connection weights between the hidden layer and output layer can be obtained through ordinary training.

In our ANN models, there are five input neurons: the radial position, the axial position, the gas pressure, the microwave power and the magnetic coil current. The numbers of the neurons in the hidden layer is 12. The output layer is our target output neuron: the plasma density. The neural networks are realised in MATLAB7.0 on the platform of Microsoft Windows XP. The neural network requires that the range of the both input values and output values should be between 0 and 1, so all data in our model are unified.

4. Results and discussions

The effects of training cycles on the network performance of two ANN models are shown in Fig. 2, in which the vertical axis indicates the difference between the training performance and the training goal. We set the training goal as 1×10^{-4} to simplify the model.

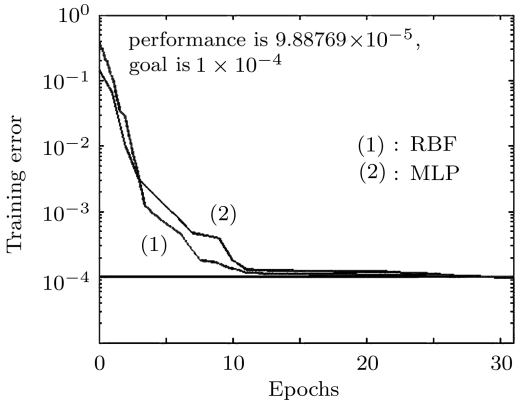


Fig. 2. The effects of the training cycles on the different network performance.

From Fig. 2 we can know that both ANN models can attain our training goal, and the RBF model is the better efficient design. In this RBF model, when the neural network was trained up to 24 times, the goal of training accuracy was attained, while the MLP model needs 30 times to attain the goal of training accuracy. It is clearly noticed from this figure 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 they gradually tend to approach the minimum value.

The accuracy of prediction in both ANN models is shown in Fig. 3. 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 horizontal axis indicates the serial numbers of those experimental data. From this figure we can conclude that the RBF model has better ability of prediction than the MLP model. It is believed that the number of neurons in the hidden layer plays an important role.

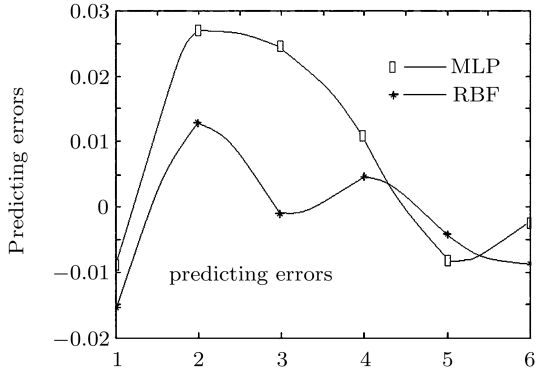


Fig. 3. Comparisons of predicting errors between two models.

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. We calculate the relative errors, the e_{MSE} and the e_{MSRE} of two models:

$$e_{\mathrm{MSE}} = \sqrt{\frac{\sum_{i=1}^{N} (V_{\mathrm{calc},i} - V_{\mathrm{meas},i})^2}{N}},$$

$$e_{\mathrm{MSRE}} = \sqrt{\sum_{i=1}^{N} \left[\frac{V_{\mathrm{calc},i} - V_{\mathrm{meas},i}}{N}\right]^2},$$

where e_{MSE} is the mean squared error, e_{MSRE} is the mean squared relative error, V_{calc} is computed value, V_{meas} is measured value, and N is the number of samples. Corresponding experimental data of test set in two ANN models are given in Table 1.

Table 1. Data used in test set and the predicting values of two ANN models.

gas pressure /Pa	microwave power/W	magnet coil current/A	axial position /cm	radial position /cm	plasma density/10^{10}cm^{-3}		
					experimental data	predicted data	
						MLP	RBF
0.06	620	143	0	8	2.260	2.126	2.211
0.11	620	143	10	4	5.406	5.529	5.473
0.13	450	153	10	14	1.985	1.863	1.919
0.11	530	148	15	8	3.824	3.854	3.896
0.08	580	158	20	10	3.405	3.365	3.352
0.15	620	162	30	2	2.284	2.189	2.168

It is shown that the predicted and experimental values are very close to each other. From Table 2, we can see that the relative errors of MLP model are between 0.7% and 6.1%, while those of RBF model are between 1.2% and 5.1%, and the e_{MSE} and the e_{MSRE} of RBF model are less than MLP model. It can be obviously concluded that the predictions of two ANN models are in good agreement with the experimental data, and the RBF model is more accurate than the MLP model. The ability of ANN model to predict the plasma density accurately in an ECR-PECVD system can be concluded.

Pankaj and Deo[21] believed that the RBF is better, which can be due to the fact that it involves fewer parameters, and does not use a large amount of user specified parameters and probably has no large requirements of sample size. We think the key is that RBF used the local-response RBF to replace the whole-response function. Because the characteristics of local-response, the RBF can perform better and faster in approach of function. In addition to this, the random centre of the RBF can avoid the choice of initial weight in the MLP. So, in our study, the RBF is more suitable than the MLP when the neuron numbers in hidden layer is the same.

Table 2. The error between experimental data and predicted data.

	MLP						RBF					
relative error	5.9%	2.3%	6.1%	0.7%	1.1%	4.2%	2.2%	1.2%	3.3%	1.8%	1.5%	5.1%
e_{MSE}	0.0996						0.0738					
e_{MSRE}	0.1660						0.1233					

5. Conclusions

In present work, a relative new technique ANN associated with nonlinear system is used in modeling the prediction of the plasma density in an ECR-PECVD plasma system. The RBF and MLP ANN models are established. The effectiveness of two ANN models are demonstrated, the results show good agreements with corresponding experimental data. The ability of ANN model to predict accurately the plasma density in an ECR-PECVD system can be concluded, and the RBF is more suitable than the MLP in our work. These indicate that ANN is a useful tool in dealing with some complex nonlinear problems of the plasma density and can obtain a considerable saving in terms of cost and time.

References

[1] Toader E I 2004 *Plasma Sources Sci. Technol.* **13** 646

[2] Chen J F and Ren Z X 1999 *Vacuum* **52** 411

[3] Mayuko K and Hiroshi M 2006 *Vacuum* **80** 771

[4] Jin X Y, Qiu X J and Zhu Z Y 2006 *Acta Phys. Sin.* **55** 5338 (in Chinese)

[5] Wang L, Cao J X, Wang Y, Niu T Y, Wang G and Zhu Y 2007 *Acta Phys. Sin.* **56** 1429 (in Chinese)

[6] Hiroshi M, Doan H T and Yoshinobu K 2005 *Surf. Coat. Technol.* **200** 850

[7] Musil J 1996 *Vacuum* **47** 145

[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] Du X L, Chen G C, Jiang D Y, Yao X Z and Zhu H S 1999 *Acta Phys. Sin.* **48** 257 (in Chinese)

[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] Wu H M, Graves D B and Porteos R K 1995 *Plasma Sources Sci. Technol.* **4** 22

[14] Sterjovski Z and Nolan D 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 J, 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

[21] Pankaj S and Deo M C 2007 *Applied Soft Computing* **7** 968